(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/29240 A2

(51) International Patent Classification[7]: C12N 15/82, 15/54, 9/12, 5/10, C07K 16/40, A01H 5/00

(21) International Application Number: PCT/NL00/00765

(22) International Filing Date: 20 October 2000 (20.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
99203480.1 22 October 1999 (22.10.1999) EP

(71) Applicant *(for all designated States except US)*: **GENETWISTER TECHNOLOGIES B.V.** [NL/NL]; Bornsesteeg 59, NL-6708 PD Wageningen (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCHMIDT, Eduard, Daniel, Leendert** [NL/NL]; Beelaertslaan 41, NL-6861 AT Oosterbeek (NL). **DE BOER, Anne, Douwe** [NL/NL]; Oude Maasdijk 32, NL-6621 AC Dreumel (NL). **VAN DER KOP, Dianne, Antoinette, Maria** [NL/NL]; Everlaan 12, NL-6705 DJ Wageningen (NL).

(74) Agent: **PRINS, A., W.**; Vereenigde, Nieuwe Parklaan 97, NL-2587 BN The Hague (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

..... *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/29240 A2

(54) Title: REGENERATION

(57) Abstract: The invention relates to the field of regeneration of cells and the vegetative propagation of (micro)-organisms or specific parts such as tissues or organs thereof, for example of those cells grown in tissue or organ culture, and more in particular to the seedless propagation of plants. The invention provides a culture method for propagation of a plant from plant starting material wherein during regeneration of said starting material, especially in the phase of the development of the shoot-root body plan, root or shoot initiation is stimulated by a recombinant gene product or functional fragment thereof, for example derived from a gene involved in the regulation of plant development allowing reducing or omitting exogenous phytohormone addition to said culture.

Title: Regeneration

The invention relates to the field of regeneration of cells, self-renewal of (micro)-organisms, the vegetative propagation of plant parts such as plant tissues or organs thereof, for example cells grown in tissue or organ culture, and more in particular to the seedless propagation of plants.

Renewal of plant and animal cells into more cells, tissues, organs and even whole plants and organisms is a process central to life that has been set to men's whims and desires already for a long time. Self-renewal of specific micro-organism starter cultures are used to ferment foods and drinks. Yet other cultures are useful for the metabolites they produce per se, such as produced by modern day's large scale fermentor cultures for the production of antibiotics or enzymes. Within the realm of animal cells, use of the renewed cultured cells, although being of fairly recent date, has taken great flight with the production of for example viral vaccines in cell- or tissue culture. Even more recent is the use of donor cells harvested from an individual, and grown and/or differentiated in culture, for transplantation purposes. Such cells (take for example bone marrow cells) are, after having been sufficiently regenerated and differentiated, proliferated or equipped with the desired characteristics, transplanted into a recipient for medical purposes. Shortly, such therapies will even include transgenic cells, transformed with modern recombinant techniques, that are thereby equipped with the desired characteristics and transplanted.

Regeneration is very well studied in plants, where it is crucial in vegetative propagation. In principle, plants can be propagated in two ways, via seeds or vegetatively without using seeds as starting material to obtain the desired plant. Both types of propagation may be impossible or undesirable under certain conditions. When propagation via seeds is unsatisfactory (when no seeds or too few of the desired seeds are formed or the desired seeds quickly loose their germination viability) then seedless propagation is often adopted. Also, when due to sexually crossing a very heterogenous progeny is or may be obtained due to its strong heterozygosity, propagation via seeds is often also considered unsatisfactory. Of course, seedless propagation of essentially seedless starting material may in a later phase give rise to the desired seeds, which can further be used to obtain the desired plants.

Within seedless propagation of plants two major fields can be distinguished: In vivo and in vitro vegetative propagation. In vivo vegetative propagation (via for example cuttings, splitting or division, layering, earthing up, grafting or budding, and other methods known to the gardener or horticulturist), has for many years played an important role in agriculture; e.g. with potatoes, apples, pears, many ornamental bulbs and tuberous plants like potatoes, many arboricultural crops, carnations, chrysanthemums, etc. Vegetative propagation is also very important in plant breeding: parent lines have to be maintained and propagated vegetatively for seed production; cloning is often required for setting up gene banks; adventitious shoot formation is needed to obtain solid mutants after mutation induction.

However, the classical methods of in vivo vegetative propagation often fall short (to slow, too difficult or too expensive) of that required or are completely impossible. In the last couple of decades, since the discovery that plants can be more rapidly cloned in vitro than in vivo, knowledge concerning vegetative propagation has grown quickly; this holds equally true for plants from temperate, subtropical as well as tropical regions. It has now even become possible to clone species by in vitro culture techniques that are impossible to clone in vivo. Different methods of in vitro vegetative or seedless propagation from plant starting material are for example using single-node cuttings, axillary branching, regeneration of adventitious organs (roots or shoots) on starting material such as explants or callus tissue and regeneration of plants from suspensions of, or even single, cells or protoplasts used as starting material. For the generation of transformed or transgenic plants, in vitro propagation is even considered a prerequisite, since it is the totipotency of individual plant cells that underlies most plant transformation systems.

To propagate plants from starting material in vitro, it is in principle necessary that at least one cell in the starting material is capable of regeneration. The ability to regenerate is for example determined by the genotype, the environmental conditions (nutrient supply, regulators and physical conditions) or the developmental stage of the plant, or combinations of these. It is well known that some families and genera have high regeneration ability: *Solanacea (Solanum, Nicotiana, Petunia, Datura,* and *Lycopersion), Crucifera (Lunaria, Brassica, Arabidopsis), Generiaceae (Achimenes, Saintpaulia, Streptocarpus) Compositae (Chicorium, Lactuca, Chrysantemum), Liliaceae*

(Lilium, Haworthia) Allium, Ornithogalum) but others, such as many decorative plants, woody species such as shrubs, conifers or trees, especially fruit trees, *Rosacea, Alstroemeria, Euphorbia,* and bulbs such as *Tulipa,* and others are notoriously difficult, even with in vitro techniques.

As indicated above, regeneration (self-renewal of (micro-)organisms and self-renewal of plants, animals or parts thereof, i.e. vegetative reproduction/propagation) can also be considered a repair strategy observed throughout the realm of micro-organisms, animal and plant species. Regeneration in plants for example comprises the formation of new tissues containing both root and shoot meristems, separate shoot or root meristems, plant organs or organ primordia from individual cells or groups of cells. Regeneration in general mimics the process of normal cellular and organ differentiation that takes place during plant development and results in the formation of the different plant organs. In normal development, early in ontogony, cells and tissues of common lineage diverge into often contrasting paths of development as they respond to developmental signals. This ability to develop in response to a specific signal is also known as cellular competence or cellular potentiality. As competent cells become committed to particular paths of differentiation, they are not readily diverted into other pathways; this restriction of the developmental potentiality of cells is referred to as determination.

Plant cells or groups of cells that under normal conditions are unable to initiate the formation of certain plant organs, meristems or organ primordia can often be stimulated by extracellular stimuli modifying the differentiation stage of the cell. Extracellular diffusible factors have shown to be essential for cellular redifferentiation in plant cells (Siegel and Verbeke, 1989 Science 244, 580-582). The perception of these signals at the cellular surface and the intracellular signal transduction that finally result in changes in transcriptional regulation provides cells with the ability to respond to such extracellular stimuli. Regeneration can result in the formation of either a shoot alone or a root alone or both together. Only after redifferentiation of a cell or tissue, regeneration is possible that results in differentiated tissue that again comprises the necessary three-dimensional layout of the emerging plant, the apical-basal or shoot-root body plan from which the mature desired plant can develop.

Indeed, central in in vitro techniques for seedless propagation are phytohormones and other factors often added to the culture medium that mimic

these extracellular stimuli. For the process of regeneration of the original starting cell into a multicellular totipotent tissue underlying and preceding somatic embryogenesis or organogenesis in vitro in cell, tissue or explant cultures which lead to a fully differentiated plant again, in general a well balanced, and per plant species often different, phytohormone addition to the culture is required. Overall, a balance is required between auxins on the one hand and cytokinin on the other. After exogenous exposure to auxin (such as 2,4-dichlorophenoxyacetic acid (2,4-D), chloramben or dicamba) or cytokinin (such as 6-benzylaminopurine or zeatine) or both, cells or tissue react by development of the shoot-root body plan, for example by forming shoots and/or roots, sometimes readily, sometimes erratically especially when the proper balance between the hormones is not properly selected.

Regeneration in vitro and especially the manipulatable nature of in vitro culture thus depends mainly on the application of these two types of hormones, and also on the ability of the tissue to respond to phytohormonal changes during culture. In general, three phases of regeneration are recognisable. In the first phase, cells in the culture acquire "competence", which is defined as the ability (not capacity) to respond to hormonal signals of organ induction. The process of acquisition of said organogenic competence is often referred to as "dedifferentiation" of differentiated cells to acquire organogenic competence. The competent cells in the culture are canalised and determined for specific tissue and organ formation for re-entry of quiescent cells into cell cycle, and organisation of cell division along the lines of the shoot-root body plan to form specific primordia and meristems under the influence of the phytohormone balance through the second phase. Especially auxin is thought to be involved in specific regenerative signal transduction pathways for adventitious root initiation, whereas cytokinin is thought to be involved in specific regenerative signal transduction pathways for adventitious shoot initiation.

Then the morphogenesis, the growing of the plant to its fully differentiated state, proceeds independently of the exogenously supplied hormones during the third phase.

Although the general principles governing regeneration via addition of exogenous phytohormones are thus fairly well understood, designing working in vitro culture protocols finding the right balance, the right time of administration or the right type or subtype of said hormones for a great many individual species

is still more or less a process of trial-and-error. However, as already indicated above, for in vitro regeneration or seedless propagation of a great many plant species is a large interest, especially for those that are in general hard to propagate.

The invention provides a culture method for propagation of a plant from plant starting material wherein, especially in the phase of the development of the shoot-root body plan, root or shoot initiation is stimulated by introducing at least one recombinant gene product or functional fragment thereof in said starting material, for example by stimulating at least one signal transduction pathway for root or shoot initiation, said gene product or gene products for example derived from a gene or genes involved in the regulation of plant development, allowing reducing or omitting exogenous phytohormone addition to said culture in the regeneration process. In a preferred embodiment the invention provides a culture method for vegetative propagation of plants from plant starting material comprising regeneration of said starting material wherein during regeneration of said starting material at least one specific signal transduction pathway for adventitious root or shoot initiation is endogenously stimulated allowing reducing or omitting exogenous phytohormone addition to said culture, in particular wherein said pathway is endogenously stimulated by a recombinant gene product derived from a gene involved in the developmental regulation of regeneration, such as a gene or gene product involved in hormone production, a gene or gene product giving feed back on hormone production, or involved in the cascade of events leading to regeneration.

Preferably, the method as provided by the invention comprises at least one step of *in vitro* culture, since it is in *in vitro* culture that the auxins or cytokinins are most widely used, in the regeneration process, especially for plants that are notoriously difficult to regenerate for vegetative propagation such as many decorative plants, woody species such as shrubs, conifers or trees, especially fruit trees, *Rosacea, Alstroemeria, Euphorbia,* and bulbs such as *Tulipa.* However, clearly, said hormones are also commonly used in *in vivo* cultures as well, (*in vivo* cultures essentially being all crop or plant culture methods traditionally used in agriculture) where such hormones are commonly added by (root or stem) dipping, spraying or watering. Especially those plants that are propagated in an essential seedless way can now be regenerated or

propagated more easily, consequently, in a preferred embodiment, the invention provides a culture method for essentially seedless propagation of plants from plant starting material comprising regeneration of said starting material wherein during regeneration at least one specific signal transduction pathway for adventitious root or shoot initiation endogenously is stimulated, e.g. by above mentioned gene product, allowing reducing or omitting exogenous phytohormone addition to said culture.

Essentially seedless propagation herein is defined in that said starting material essentially comprises no seeds, or at least that seed possibly present in said starting material does not lay at the basis of the regeneration of said starting material or does not develop into the desired plant. However, as one aspect of the culture method comprising regeneration as provided by the invention, during or after the process of regeneration or propagation according to the invention seed may be formed, from which even a desired plant may develop, which is a result of the propagation according to the invention, rather than that it lays at the basis thereof.

In particular, the invention provides a culture method wherein said starting material comprises an individual plant cell or protoplast or explant or plant tissue, materials which are commonly used in in vitro culture methods whereby the addition of phytohormones was thought to be axiomatic. Now such addition is no longer necessary or can be reduced, providing an easier way of in vitro culture, wherein not such an intricate balance between the addition of the various hormones has to be sought.

The invention provides manipulation of propagation characteristics of for example plant tissue. Numerous plant species are propagated in tissue culture in order to obtain large amounts in a relative short period of time. Using the invention it is relatively easy to increase the multiplication factor several times. For several notoriously difficult species, like shrubs, trees en various bulbous species it is now also possible to use esssentially seedless propagation, and especially in vitro culture, when using the invention. The regeneration capacity of cells or tissue isolated from these plants is increased significantly, thereby increasing the multiplication factor by introducing of certain bioactive molecules, like nucleic acid or (modified) protein. The nucleic acids or proteins may be introduced by the methods known in art, like particle gun bombardment, electroporation, micro-injection or other techniques described in the introduction.

The introduced molecules are either nucleic acid, being RNA, or naked DNA with a small chance of becoming integrated in the genome, or (modified) protein product. The molecules will in general be lost during the regeneration process and are therefore only transiently present. The nucleic acids that may be used encode or produce proteins that stimulate the regeneration process and reduce or eliminate the use of exogenously added planthormones. The proteins that may be added are the protein products of these nucleic acids or their modified forms. Examples of molecules with the above described characteristics are proteins or genes coding for proteins involved in the regulation of plant development or perception of plant hormones. By using the invention the multiplication factor can be increased so much that it will be possible to use in vitro propagation techniques in a broader sense and also for the more difficult species, Also, by using the invention it is relatively easy to permanently increase the propagation characteristics for these plants. The regeneration capacity of these plants can be increased significantly if these plants are made transgenic by introducing a gene coding for proteins involved in the regulation of plant development or perception of plant hormones or more specific a gene coding for a product stimulating or inducing one signal transduction pathway for root or shoot initiation or even more specific a gene coding for a representative of the plant receptor kinase family RKS. Transformation can be achieved using the techniques known in the field like Agrobacterium mediated transformation, particle gun bombardment, the above described marker-free transformation system or others and select for non-lethal expressors of the gene.

In one preferred embodiment, the invention provides a culture method according to the invention wherein said starting material comprises a desired somatic mutation. Mutations can occur in any cell of a living organism, but are only transferred to the offspring when this mutation occurred in those cells from which gametophytic cells of that organism are derived. Somatic mutations are usually lost unless the tissue in which the mutation is apparent is vegetatively propagated or if cells in this tissue are regenerated to form an intact new organism. Using the technology described in this invention the rescue of somatic mutations in plants is provided. Somatic, but also generative tissue is stimulated to regenerate by the introduction of bioactive molecules, like nucleic acid or (modified) protein as provided by the invention. The nucleic acids or proteins may be introduced by the methods known in art, like particle gun bombardment,

electroporation, micro-injection or other techniques described. The introduced molecules are either nucleic acid, being RNA, or naked DNA with a (not necessarily) small chance of becoming integrated in the genome, or (modified) protein product. The molecules will in general be lost during the regeneration process and are therefore in general only transiently present. The nucleic acids that may be used encode proteins that stimulate the regeneration process and reduce or eliminate the use of exogenously added planthormones. The proteins that may be added are the protein products of these nucleic acids or their modified forms. Examples of molecules with the above described characteristics are proteins or genes coding for proteins involved in the regulation of plant development or perception of plant hormones. Alternatively somatic mutations may have been created by treatment of seeds with mutagenic agents, like colchicines, EMS, radiation or carcinogenic substances etc. The sectors in these mosaic plants grown from these treated seeds will be screened for desirable phenotypes. The interesting sectors will subsequently be isolated and used as starting material for regeneration by the above-described invention in order to obtain clonal propagation of these desired traits.

In another preferred embodiment, the invention provides a culture method according to the invention wherein said starting material comprises transgenic material. These days transgenic plants are being produced rapidly, albeit often in only limited numbers. To rapidly acquire sufficient numbers of plants for further propagation under field conditions, in vitro culture techniques are widely used. The invention now provides a method wherein little or no attention has to be given to phytohormone levels in such transgenic plants cultures.

In particular, the invention provided a method wherein said starting material additionally comprises starting material comprising a recombinant nucleic acid encoding a desired trait. The invention herewith provides essentially marker-free transformation, or at least it provides plants that after transformation and propagation are essentially marker-free. A recombinant nucleic acid encoding a desired trait, that one would like to integrate in a plant's genome is provided to at least part of said starting material with gene delivery vehicles or methods, such as vectors, particle bombardment, electroporation, micro-injection or other techniques described in the art. Cells comprising said recombinant nucleic acid are also provided according to the invention with at

least one recombinant gene product or functional fragment thereof, for example by stimulating at least one signal transduction pathway for root or shoot initiation, said gene product or gene products for example derived from a gene or genes involved in the regulation of plant development, allowing reducing or omitting exogenous phytohormone addition to said culture. In particular, the invention provides a culture method for vegetative propagation of plants from plant starting material having been provided with a recombinant nucleic acid encoding a desired trait comprising regeneration of said starting material wherein during regeneration of said starting material at least one specific signal transduction pathway for adventitious root or shoot initiation is endogenously stimulated allowing reducing or omitting exogenous phytohormone addition to said culture, in particular wherein said pathway is endogenously stimulated by a recombinant gene product derived from a gene involved in the developmental regulation of regeneration, such as a gene or gene product involved in hormone production, a gene or gene product giving feed back on hormone production, or involved in the cascade of events leading to regeneration.

In a preferred embodiment, said recombinant nucleic acid encoding a desired trait has additionally been provided with means for nuclear targeting and/or integration in a plant genome. Such means can be nucleic acid signals incorporated with the recombinant nucleic acid encoding the desired trait, or proteinaceous substances such as transposases, or viral or bacterial proteins (such as Vir-proteins) to protect the recombinant nucleic acid inside the cell, taking care of proper targeting towards the nucleus and/or stimulating proper integration.

Even more preferred, the invention provides a method wherein said starting material comprises a to be transformed individual plant cell or protoplast or explant or plant tissue comprising recombinant nucleic acid encoding a desired trait among other, non-transformed starting material from which the transformed material has to be selected.

In general, as a part of the process of for example plant transformation, dominant selectable markers are used to select transgenic cells from which transgenic plants can be regenerated. For one thing, these marker genes are generally superfluous once an intact transgenic plant has been established. Furthermore, selectable marker genes conferring for example antibiotic or herbicide resistance, used to introduce economically valuable genes into crop

plants have major problems: detoxification of the selective agent by expression of a modifying enzyme can enable untransformed cells to escape, dying untransformed cells release products which are toxic and inhibit the regeneration of transformed cells, the selective agents may have negative effects on proliferation and differentiation of cells, there is uncertainty regarding the environmental impact of many selectable genes, and it is difficult to perform recurrent transformations using the same selectable marker to pyramid desirable genes. The invention now provides a method reducing or omitting selective agent addition to said culture.

Attempts have been made earlier to design transformation systems allowing marker gene elimination to obtain marker-free transformants of diverse plant species whereby the marker gene used is removed from the transformed cell after it has performed its task. One method involves co-transformation of cells mediated by Agrobacterium tumefaciens with binary vectors carrying two separate T-DNAs, one for example comprising a drug-resistance selection marker gene, another comprising the desired gene, followed by conventional out-breeding the undesired drug-resistance gene, that is thought to localise at a different locus than the desired gene. Although drug sensitive transformants comprising the desired gene may be thus obtained it is not clear whether all these transformants are indeed totally free of (non or partly functional) selection marker-gene or fragments thereof. Also, the selective agent initially used still has the unwanted negative effects on proliferation and differentiation of plant cell during the transformation process. Furthermore, the method requires sexual crossing which limits it to plant species where sexual crossing, and not vegetative reproduction, is the practical method of reproduction, and practically limits it even further to those plant species with a sufficient short generation time.

One strategy currently available to eliminate the superfluous marker after the cell has been transformed without the need to sexually cross plants is the MAT vector system. However, said system relies on intrinsic post-transformational excision of the selection gene which is comprised in a transposable element, an event which only haphazardly occurs and reduces the final efficiency of the transformation process.

Yet another strategy involves site specific recombination such as seen with the Cre-Lox system whereby in a first transformation the selection-marker

gene is inserted at a previously determined specific site, allowing selection of transformed cells, after which in a second transformation comprising the introduction of a site specific recombinase, the selection-marker gene is again excised from the genome.

Needles to say that, apart from other problems, the prerequisite of having a suitable site in the to be transformed cell available restricts said method to those organisms of which the genome is well known. The invention now provides a method to obtain transformed plants by in vitro culture wherein said transgenic material is devoid of a selectable marker gene conferring resistance to an selective agent. Resistance to selective agents is no longer needed since according to the invention the transformed material is equipped with the necessary recombinant gene product or gene products or functional fragment(s) thereof derived from a gene involved in the regulation of plant development allowing reducing or omitting exogenous phytohormone addition to said culture, thereby giving preferred growth conditions to the transformed cells over those non-transformed cells that have not been provided with said gene product or functional fragment thereof. In particular, the invention provides a culture method for vegetative propagation of plants from transformed plant starting material comprising regeneration of said starting material wherein during regeneration of said transformed starting material at least one specific signal transduction pathway for adventitious root or shoot initiation is endogenously stimulated allowing reducing or omitting exogenous phytohormone addition to said culture, in particular wherein said pathway is endogenously stimulated by a recombinant gene product derived from a gene involved in the developmental regulation of regeneration. The beauty of it is that no selectable marker gene conferring resistance to a selective agent has to be introduced in said material at all, thereby obviating the need to deplete the transformed material of such marker genes afterwards. In particular, the invention thus does not make use of resistance to antibiotic or herbicides, and does nor carry all the disadvantages associated herewith.

In short, most plant transformation systems are based on the selection for herbicide or antibiotic resistance or selection for transformants is based on the presence of an additional selection marker besides the trait itself. Using the technology described in this invention, markerless transformation in plants is provided. This new transformation/regeneration (t/r) system for example consist

of two components (Fig. 20). A first component in this example is the trait, which may be present between the borders of Agrobacterial T-DNA, but apart from a suitable promoter no other DNA is needed. This first component may be single or double stranded DNA and may be *in vitro* coated with the VirE2 protein and/or a molecule of VirD2 (preferentially covalently attached to the 5'-end of this DNA). The Vir-proteins may be present to protect the DNA inside the plant cell, take care of proper targeting towards the nucleus and will stimulate proper integration into plant DNA. Tissue will be stimulated to regenerate by the introduction of certain bioactive molecules. These bioactive molecules act as the second component. The second component is either nucleic acid, being RNA, or naked DNA with a small chance of becoming integrated in the genome, or (modified) protein product.

The nucleic acids or proteins (second component) may be introduced mixed with the first component by the methods known in art, like particle gun bombardment, electroporation, micro-injection or other techniques described in the introduction. Both components have to be present in the plant cell together in sufficient quantities, but the ratio between the two components may vary depending on the species and the preferred number of integration's of the trait in the plant DNA. The second component will preferably be lost during the regeneration process and is therefore only transiently present, whereas the first component has a high change of becoming integrated into the plant genome. The second component is a nucleic acid or a mixture of nucleic acids that will produce proteins that stimulate the regeneration process and reduce or eliminate the use of exogenously added planthormones or is the protein product or a mixture of products of these nucleic acids or their modified forms or a mixture of both. Examples of molecules with the above described characteristics are proteins, or genes coding for proteins involved in the regulation of plant development or perception of plant hormones. The main advantages of the this t/r-system are, as explained with the example of figure 20:

- only the trait is introduced into the plant DNA; apart from the T-DNA borders (Only in the case when VIR proteins are used, it is necessary to include T-DNA borders onto the trait DNA), if present, no other unwanted DNA, like a selection marker, is present. In order to allow the process of homologous recombination of the trait DNA into the corresponding endogenous DNA on the plant genome, genes or gene

products encoding At R51, AtRAD51 or RecA or gene products with similar function can be applied in the second component in order to result in transient expression of the recombinase. After targeting and localized integration of the trait DNA, the recombinase is lost.

- the principle of regeneration is universally applicable
- the amount of exogenous plant hormones for regeneration can be reduced or omitted

active selection is not necessary as mainly transformed cells will regenerate.

Said gene involved in the regulation of plant development can be selected from a great many genes already known, or yet to be determined, to be involved in regeneration. Examples of such genes are clavata (Clark et al., 1997, Cell 89, 575-585) and primordia timing genes (Mordhorst et al, 1998 Genetics 149, 549-563), which are stimulating regeneration when inactivated, Leafy-Cotelydon gene (LEC, Lotan et al., 1998, Cell 93, 1195-1205), the KAPP gene (Stone et al., 1994, Science 266, 793-795; Stone et al., 1998, Plant Physiol. 117, 1217-1225), IPT (Morris, R.O., 1986 Annu. Rev. Plant Physiol. 37, 509-538), WUSCHEL (Mayer et al. 1998 Cell 95, 805-815; Schoof et al. 2000 Cell 100, 635-644), KNAT1&2 (the Arabidopsis kn1-like gene) (Chuck et al. 1996. Plant Cell 8, 1277-1289; Lincoln et al. 1994 The Plant Cell 6, 1859-1876), SHOOT MERISTEMLESS gene (Endrizzi et al. 1996 Plant J. 10, 967-979), CUP-SHAPED COTYLEDON (Aida et al. 1999 Development 126, 1563-1570), CYCLIN D (Cockcroft et al. 2000 Nature 405, 575-579; Riou-Khamlichi et al. 1999 Science 283, 1541-1544),
CKI1 (Kakimoto 1996 Science 274, 982-985), AINTEGUMENTA (Mizukami and Fischer 2000 PNAS 97, 942-947; Krizek 1999 Dev. Genetics 25, 224-236), SBP-box proteins (Cardon et al. 1999 Gene 237, 91-104), CDC2a (Hemerly et al. 1993 The Plant Cell 5, 1711-1723), which are genes that stimulate regeneration when induced or overexpressed, or antagonists thereof or others that are involved in the regulation of plant development in the broadest sense, such as can be found by studying plant embryogenesis or organogenesis on the molecular level. In particular, a population of gene products involved in regeneration is represented by the intracellular signal transduction factors that are directly phosphorylated by RKS protein and thereby activated.

In a preferred embodiment, the invention provides a method according to the invention wherein said gene involved in the regulation of plant development

encodes a leucine-rich repeat containing receptor-like kinase, such as present in plant database collections, with homology to the extracellular domain of the Arabidopsis RKS protein family, such as:

GB:AW011134 AW011134 ST17B03 Pinus taeda

GB:LELRPGENE X95269 L.esculentum

GB:AI775448 AI775448 EST256548 Lycopersicon esculentum

GB:AI496325 AI496325 sb05c09.y1 Gm-c1004 Glycine

GB:AI487272 AI487272 EST245594 Lycopersicon esculentum

GB:AI441759 AI441759 sa82d08.y1 Gm-c1004 Glycine max

GB:AI782010 AI782010 EST262889 Lycopersicon esculentum

GB:AI772079 AI772079 EST253179 Lycopersicon esculentum

GB:SBU62279 U62279 Sorghum bicolor

GB:C22645 C22645 C22645 Oryza sativa

GB:D49016 D49016 RICS15625A Oryza sativa

GB:AI776399 AI776399 EST257499 Lycopersicon esculentum

GB:AI776208 AI776208 EST257308 Lycopersicon esculentum

GB:AI352795 AI352795 MB61-10D PZ204.BNlib Brassica napus

GB:AQ578072 AQ578072 nbxb0092C18f Oryza sativa

GB:C95313 C95313 C95313 Citrus unshiu Miyagawa

GB:AI162893 AI162893 A026P38U Hybrid aspen

GB:AI782076 AI782076 EST262955 Lycopersicon esculentum

GB:AI726177 AI726177 BNLGHi5165 Cotton

GB:AI777982 AI777982 EST258861 Lycopersicon esculentum

GB:AI774881 AI774881 EST255981 Lycopersicon esculentum

GB:AI896737 AI896737 EST266180 Lycopersicon esculentum

GB:AI676939 AI676939 605047A07.x1 Zea mays

GB:D40598 D40598 RICS2674A Oryza sativa

GB:OSU82168 U82168 Oryza sativa

GB:SBRLK1 Y14600 Sorghum bicolor

GB:AI495359 AI495359 sa97a09.y1 Gm-c1004 Glycine max

GB:C96041 C96041 C96041 Marchantia polymorpha,

or such as present in plant database collections, with homology to the intracellular domain of the Arabidopsis RKS protein family, such as:

GB:AI896277 AI896277 EST265720 Lycopersicon esculentum

GB:AU056335 AU056335 AU056335 Oryza sativa
GB:AA738546 AA738546 SbRLK4 Sorghum bicolor
GB:AA738544 AA738544 SbRLK2 Sorghum bicolor
GB:AA738545 AA738545 SbRLK3 Sorghum bicolor
GB:SBRLK1 Y14600 Sorghum bicolor
GB:AI729090 AI729090 Gossypium hirsutum
GB:AI920205 AI920205 Pinus taeda
GB:AI896183 AI896183 EST265626 Lycopersicon esculentum
GB:AI967314 AI967314 Lotus japonicus
GB:AI730535 AI730535 BNLGHi7007 Gossypium hirsutum
GB:AF078082 AF078082 Phaseolus vulgaris
GB:CRPK1 Z73295 C.roseus
GB:C22536 C22536 C22536 Oryza sativa
GB:C22530 C22530 C22530 Oryza sativa
GB:ZMA010166 AJ010166 Zea mays mRNA
GB:AQ271213 AQ271213 Oryza sativa,

or known from Schmidt et al (1997, Development 124, 2049-2062, WO 97/43427), where for example stable transformation, but not regeneration, nor transient expression nor use in selection, of plants with SERK (RKS0) is considered. Also applicable in a method according to the invention are bacterial genes or fragments thereof such as the AK-6b gene (Wabiko et al, Plant Physiol. 1996, 939-951) or the rolABC genes (Jasik J, Plant Science, 1997, 57-68), however, where only regeneration by stable transformation is intended, plant genes such as those disclosed herein are preferred.

In a preferred embodiment, the invention provides a method according to the invention wherein said gene involved in the regulation of plant development encodes a leucine-rich repeat containing receptor-like kinase, wherein said receptor-like kinase is a representative of a plant receptor kinase family RKS such as shown in figure 3.

In particular, the invention provides a method wherein said gene product or functional fragment thereof is derived from a receptor-like kinase that comprises an N-terminal signal sequence, an extracellular region comprising a leucine zipper domain, a disulphate bridge domain, a leucine rich repeat domain comprising 3-5 leucine rich repeats, a transmembrane domain, an intracellular

region comprising an anchor domain, a serine/threonine kinase domain and/or a C-terminal leucine rich repeat domain.

These genes encode membrane spanning proteins having a particular function in signal transduction, thereby being prime candidate genes to provide gene products or functional fragments thereof to be employed in a method of the current invention.

In particular, the invention provides a method wherein said receptor-like kinase is encoded by a nucleic acid which in *Arabidopsis thaliana* comprises a sequence as shown in anyone of figures 4 or 8 to 20. Suitable receptor kinase-like genes from plants other than *Arabidopsis thaliana*, such as *Daucus carota, Rosa, Gerbera, Chrysanthemum, Alstroumeria, Lilium, Tulipa, Dyanthus, Cymbidium, Gypsopays, Ficus, Calangoe, Begonia, Phalasnopsis, Rhonondendrum, Spatiphilus, Cucubitaceae, Solanaceae,* and grasses such as cereals are easily found using the *Arabidopsis thaliana* sequences provided herein by methods known in the art. In general for each RKS gene identified in *Arabidopsis thaliana* a corresponding RKS gene is present in individual species of both monocotyledon as well as in dicotyledon plants. The invention provides a method wherein said receptor-like kinase is encoded by a plant derived nucleic acid corresponding or homologous to a nucleic acid which in *Arabidopsis thaliana* comprises a sequence as shown in anyone of figures 4 or 8 to 20. Corresponding or homologous RKS genes and gene products in plant species other than Arabidopsis thaliana are isolated by various approaches. For example by screening of cDNA and genomic libraries using Arabidopsis RKS cDNA probes under low stringency hybridisation/washing conditions as described above, alternatively by the use of degenerated RKS primers (for example primer combination RKS B forward and RKS E reverse as shown herein in order to amplify an exon fragment of the desired gene. Full length cDNA clones can further be obtained by race and tail PCR approaches. Also, the generation of antibodies recognising conserved or distinct and specific regions within different members of RKS gene family within a plant species allow the desired isolation. Alternatively, specific antibodies are generated that recognise one specific RKS gene product in a variety of plant species. These antibodies are used to screen cDNA expression libraries of plant species. Furthermore, it is possible to screen for RKS-homologous sequences in electronic databases. Searches are performed both on nucleotide and on amino acid level. Additionally, RKS genes and gene

products in plant species other than Arabidopsis thaliana are isolated for example by two or three hybrid screenings in yeast with RKS clones in other to isolate (hetero) dimerizing members of this RKS family in similar or unrelated plant species.

In one embodiment, the invention provides a method for propagation of a plant from plant starting material wherein during regeneration of said starting material at least one signal transduction pathway for root or shoot initiation is stimulated by a recombinant gene product or functional fragment thereof derived from a gene involved in the regulation of plant development allowing reducing or omitting exogenous phytohormone addition to said culture, wherein said gene product or functional fragment thereof is introduced in at least a part of the starting material by transformation. The invention also provides the introduction of regenerating gene constructs into cells which can lead to the regeneration of the cell itself or to the induction of regeneration processes in neighbouring cells, even somatic embryos resulting from said induced cells are provided herewith. Individual transformed cells are generated that are essential for the differentiation state of surrounding cells. Introduction of such an inducing regenerator as provided herewith into plant cells results in the formation of a proliferation of neighbouring cells and the formation of new plants or parts thereof from these proliferating cell masses. The originally transformed plant is not necessarily included in the proliferation process itself an is therefore not necessarily part in the resulting regenerating plants or parts thereof. This specific from of induced regeneration of neighbouring cells provide herewith gives the option to regenerate plants that do not contain the introduced gene or gene product, and therefore represents a method to induce regeneration without the necessity to introduce gene products into an originating cell population and having to maintain these gene products or nucleic acids encoding therefore. An example of the process of induced induction is shown in Figure 6F, where a single GUS positive cell marks the original introduction site for the bombarded DNA constructs. Above this cell, a proliferating cell mass has been formed that is clearly GUS negative. On top of this induced proliferated cell mass, we could detect several structures that morphologically represent somatic embryos. These somatic embryos develop from the borders of the proliferating cell mass as previously described (Schmidt et al. 1997, Development 124, 12049-2062). Somatic embryos provide an excellent source of regenerating plant since all the

organs and plant parts are formed by similar processes as take place during zygotic embryogenesis. This observation clearly indicates the potential of this class of regenerating molecules to induce a proliferating, non-transformed cell mass from which new plantlets can be regenerated. It provides the means to induce somatic embryos directly on living plant tissues, even without the prior need to introduce an in vitro culture procedure.

Again, transformation as provided here can be thus either in a stable fashion where the introduced genetic information or nucleic acid is integrated into the nuclear, chloroplast or mitochondrial genome, and is either constitutively or inducibly expressed but preferably is transient, wherein the nucleic acid is not introduced into the genome and gets lost after a certain period after introduction. Transformation of recombinant DNA or RNA into the cell or protoplast can take place in various ways using protocols known in the art, such as by particle bombardment, micro-injection, Agrobacterium-mediated transformation, viral-mediated transformation, bacterial conjugation, electroporation, osmotic shock, vesicle transport or by direct gene transfer, with or without the addition of a proteinaceous substance bound to the nucleic acid molecule. Integration of a proteinaceous substance into cells or protoplast can be facilitated along the lines of the transformation protocols as described above. A cell or protoplast thus having been provided with a gene product (i.e. a DNA, RNA or proteinaceous substance or functional fragment thereof) derived from a gene involved in the regulation of plant development can now regenerate on its own, allowing reducing or omitting exogenous phytohormone addition to the culture that comprises that cell or protoplast. The process of vegetative propagation is hereby very much simplified, large numbers of plants with an identical genetic background can now be obtained staring from starting material with the desired characteristics.

In a preferred embodiment, the present invention provides a method for propagation of a plant from plant starting material wherein said starting material comprises a cell or protoplast transformed with a desired nucleic acid sequence intended to provide the resulting transgenic plant arising from that cell or protoplast with desirable characteristics. Such a cell or protoplast, according to the invention having been provided with a gene product (i.e. a DNA, RNA or proteinaceous substance or functional fragment thereof), for example derived from a gene involved in the regulation of plant development can now regenerate

on its own, allowing reducing or omitting exogenous phytohormone addition to the culture that comprises that transformed cell or protoplast. Selection for regenerating cells or tissues after the transformation of the desired sequence together with the regenerating gene product results in the recovery of only those plants or plant material that contain the desired nucleic acid sequence, preferably integrated in a stable fashion in the plant's genome, and the regenerating gene product, thereby providing a selection of the desired transgenic plant based on the selective regeneration of the transformed starting material.

In a preferred embodiment, the invention provides a method wherein the regenerating gene product is only transiently expressed, wherein the regenerating gene product or its coding sequence is not introduced into the genome and gets lost after a certain period after introduction, hereby providing an essentially marker-free transgenic plant as end-product, containing only the desired transgenic nucleic acid, and not the nucleic acid encoding the selection marker used: the regenerating gene product.

Furthermore, the invention provides plant or plant material obtainable by a method according to the invention, propagated along the lines or using a method herein disclosed. In particular, the invention provides a plant or plant material obtainable by in vitro vegetative or seedless propagation according to the invention from plant starting material, for example using single-node cuttings, axillary branching, regeneration of adventitious organs (roots or shoots), or starting material such as explants or callus tissue or suspensions of, or even single, cells or protoplasts, in particular wherein said starting material comprises transgenic material, said transgenic plant or plant material according to the invention preferably being free of a selection marker gene.

The invention furthermore provides an isolated and/or recombinant nucleic acid encoding a receptor-like kinase or a functional fragment or functional equivalent thereof, corresponding to or capable of hybridising to a nucleic acid molecule as shown in anyone of figures 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20, or its complementary nucleic acid. Such a nucleic is obtained as described above. In a preferred embodiment, such a nucleic acid is at least 75% homologous, preferably at least 85%, more preferably at least 90%, or most preferably at least 95 % homologous to a nucleic acid molecule or to a functional equivalent or functional fragment thereof, as shown in anyone of figures 8, 9, 10,

11, 12, 13, 14, 15, 16, 17, 18, 19 or 20, or its complementary nucleic acid, for example derived from *Arabidopsis thaliana*.

Also, the invention provides a vector comprising a nucleic acid according to the invention. Such a vector is preferably capably of providing stably or transient transformation of a cell by providing said cell with nucleic acid (DNA or RNA) or protein derived from a nucleic acid according to the invention. A variety of methods to provide cells with nucleic acid or protein are known, such as electroporation, liposome-mediated transfer, micro-injection, particle gun bombardment or bacteria-mediated transfer. RNA can for example be produced in vitro from appropriate vector constructs incorporating sites such as SP6, T7 or T3. Protein is produced in vitro in for example yeast or bacterial or insect cells, or other appropriate cells known in the art. DNA can be delivered as linear or circular DNA, possibly placed in a suitable vector for propagation.

Furthermore, the invention provides a host cell comprising a nucleic acid or a vector according to the invention. In a preferred embodiment, such a host cell is a transformed cell additionally comprising a desired, but most times totally unrelated, nucleic acid sequence, preferably integrated in a stable fashion in its genome. Even more preferred is a host cell according to the invention wherein the nucleic acid or vector according to the invention is only transiently expressed. Of course it is preferred to use a nucleic acid, vector or host cell according to the invention for use in a culture method as provided by the invention. The invention also provides a method for determining a developmental stage of a plant comprising detecting in said plant or parts thereof a nucleic acid or a proteinaceous substance according to the invention. Said detection is thus aimed at using receptor kinase genes or gene products belonging to the RKS family, or fragments thereof, as markers for plant development.

The invention furthermore provides an isolated or recombinant proteinaceous substance comprising an amino acid sequence as shown in anyone of figures 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20, or a functional equivalent or functional fragment thereof. Proteinaceous substance herein is defined as a substance comprising a peptide, polypeptide or protein, optionally having been modified by for example glycosylation, myristilation, phosporylation, the addition of lipids, by homologous or heterologous di-or multimerisation, or any other (posttranslational) modifications known in the art.

Based on sequence composition, the N-terminal domain of predicted amino acid sequences of the RKS gene family represents a signal peptide, indicating that this region of the protein is extracellular. The length of this signal sequence and the predicted cleavage sites have been established using a prediction program: http://genome.cbs.dtu.dk/services/SignalP/. This domain is followed by a short domain containing a number of leucine residues, seperated from each other by 7 amino acid residues. Based on the conservation of these leucines in an amphipathic helix , this domain represents a leucine zipper domain that mediates protein dimerization through formation of a short coiled-coil structure (Landschultz WH, Johnson PF, and McKnight sSL (1988) Science 240, 1759-1764). In RKS proteins, this leucine zipper domain is likely to be involved in receptor hetero/homo dimerization. The next domain contains 2 conserved cysteine residues that forms a disulphate bridge. The subsequent domain represents a leucine rich repeat (LRR) region with 3-5 LRRs of approximately 24 amino acids each. In animals, this domain is known to be involved in protein-protein interactions (Kobe B and Deisenhofer J (1994) TIBS 19, 415-420). In plants the extracellular LRR region is predicted to be necessary for ligand and elicitor binding. At the C-terminal part of the LRR region of most RKS proteins, another conserved couple of cysteine residues is involved in the formation of another disulphate bridge. At both ends, the LRR domain is thus surrounded by two disulphate bridges. The next domain contains a relatively high number of P and S amino acid residues, and shows similarity with cell wall proteins like extensins. Prediction server programs like http://genome.cbs.dtu.dk/services/NetOGlyc/ indicate the presence of multiple O-glycosylation sites within this domain. This domain might have similar functions as extensins and provide interaction sites with multiple cell wall components, thus forming a stable immobilised interaction with the cell wall in which the complete extracellular region of RKS proteins is embedded. The next domain represents a single transmembrane helical domain, as predicted by the program http://genome.cbs.dtu.dk/services/TMHMM-1.0/. The end of this domain, and the beginning of the intracellular cytoplasmic domain, contains a small number of basic K and R residues. The next domain is relatively acidic. The next large domain shows extensive homology with the family of plant serine, threonine receptor kinases. Autophosporylation studies on SERK (Schmidt et al. 1997) have shown that this domain shows serine, threonine kinase activity. Within the

kinase domain, several RKS proteins like RKS0 and RKS8 contain a putative 14-3-3 binding site represented by the core sequence RxpSxP, in which x represents any amino acid (Yaffe MB, Rittinger K, Volinia S, Caron PR, Aitken A, Leffers H, Gamblin SJ, Smerdon SJ and Cantley LC (1997) Cell 91, 961-971). (Auto)phosphorylation of the S residue within this sequence as a result of ligand-mediated receptor-kinase activation would thus allow the binding and subsequent activation of 14-3-3 proteins. The next domain has an unknown function although the conservation of WD pair residues suggests a function of a docking site for other proteins. The C-terminal intracellular domain contains again part of a single LRR sequence, and might therefore be involved in protein-protein interactions. Preferably such a proteinaceous substance according to the invention is encoded by a nucleic acid according to the invention or produced by a host cell according to the invention.

In particular, the invention provides a proteinaceous substance for use in a culture method according to the invention. Introduction of a proteinaceous substance into cells or protoplast can be facilitated along the lines of the transformation protocols as known in the art. A variety of methods are known, such as micro-injection, particle gun bombardment or bacteria-mediated transfer. A cell or protoplast thus having been provided with a proteinaceous substance or functional fragment thereof derived from a gene involved in the regulation of plant development can now regenerate on its own, allowing reducing or omitting exogenous phytohormone addition to the culture that comprises that cell or protoplast. The process of vegetative propagation is hereby very much simplified, large numbers of plants with an identical genetic background can now be obtained staring from starting material with the desired characteristics. Proteins or peptides, encoded for by the RKS genes, are produced by expressing the corresponding cDNA sequences, or parts thereof *in vitro* or in an *in vivo* expression system in E.coli yeast, Baculovirus or animal cell cultures. The expressed protein sequences are purified using affinity column purification using recombinant Tag sequences attached to the proteins like (HIS)6 tags. Tags are removed after purification by proteolytic cleavage. The resulting protein sequence encodes a functionally active receptor-kinase, or a derivative thereof. In a preferred embodiment, the protein contains a (constitutive) active kinase domain. The purified recombinant protein is introduced into plant cells in order to induce regeneration from these cells in a transient fashion. Proteins are

introduced by methods similar as described for the introduction of nucleotide sequences, such as liposome-mediated transfer, micro-injection, electroporation, particle gun bombardment or bacteria-mediated transfer. If so desired, modification of recombinant proteins like glycosylation, disulphate bridge formation, phosphorylation etc. can be optimized in order to obtain an optimal efficiency in protein stability and activity.

Also, the invention provides an isolated or synthetic antibody specifically recognising a proteinaceous substance according to the invention. Such an antibody is for example obtainable by immunising an experimental animal with a proteinaceous substance according to the invention or an immunogenic fragment or equivalent thereof and harvesting polyclonal antibodies from said immunised animal, or obtainable by other methods known in the art such as by producing monoclonal antibodies, or (single chain) antibodies or binding proteins expressed from recombinant nucleic acid derived from a nucleic acid library, for example obtainable via phage display techniques. Such an antibody can advantageously be used in a culture method according to the invention, for example to identify cells comprising a regenerating gene product as identified above. With such an antibody, the invention also provides a proteinaceous substance specifically recognisable by such an antibody according to the invention, for example obtainable via immunoprecipitation, Western Blotting, or other immunological techniques known in the art. Also, the generation of such antibodies recognising conserved or distinct and specific regions within different members of RKS gene family within a plant species allow the desired isolation of RKS-homologues or recognise a specific RKS gene product in a variety of plant species. These antibodies are also used to screen cDNA expression libraries of plant species to screen for RKS-homologues. The invention, and use as provided of a nucleic acid, a vector, a host cell, a proteinaceous substance or an antibody according to the invention in a method according to the invention is further explained in the detailed description without limiting the invention.

Detailed description.

In order to isolate genes involved in the developmental regulation of regeneration in plants, the different members of a family of genes were identified of which the expression was present in developing influorescenses. Within this

tissue a large number of different organ primordia are initiated from the influorescence meristems. As a model plant species Arabidopsis thaliana was choosen, based on the presence of many well characterized genetic mutations and the availability of genetic information in databases.

The differentiation stage is highly stable in vivo, yet in response to nuclear transplantation or cell fusion, the nuclei of differentiated cells exhibit a remarkable capacity to change, both in animal and in plant cells (Blau, 1989). The ability to change the differentiation stage provides cells and tissues with the ability to adapt towards their environment. Normally only a small number of stem cells have the ability to differentiate into different cell types. In plants, the only cells that are truly totipotent are the zygotes, consisting of fused egg cells and sperm. From these dipoid totipotent cells all other differentiated cell types are derived.

Regeneration is a vegetative reproduction or repair strategy observed in a large number of animal and plant species. Regeneration in plants is defined as the formation of new tissues containing both root and shoot meristems, separate shoot or root meristems, plant organs or organ primordia from individual cells or groups of cells. Regeneration mimics the process of normal cellular and organ differentiation that takes place during plant development and results in the formation of the different plant organs. However, plant cells or groups of cells that under normal conditions are unable to initiate the formation of certain plant organs, meristems or organ primordia can be stimulated by either extracellular stimuli or intracellular modification of the differentiation stage of the cell. Regeneration can take place under either in vivo or in vitro conditions.

Regeneration does not include the process of apomixis, wherein specific forms of vegetative plant reproduction are taking place in seeds. Extracellular diffusible factors have shown to be essential for cellular redifferentiation in plant cells (Siegel and Verbeke, 1989). The perception of these signals at the cellular surface and the intracellular signal transduction that finally result in changes in transcriptional regulation provides cells with the ability to respond to such extracellular stimuli.

In a search for gene products with the ability to regulate cellular differentiation we concentrated on genes involved in perception and transmission of intercellular differentiation signalling. Extracellular signals in animal cells are normally perceived by an high affinity binding compound, the sensor molecule.

Extracellular signalling factors are further referred to as ligands and their cellular binding partners are defined as receptors. Upon binding, the extracellular signal can result in modification of the receptor, resulting in transmission of the signal over the cellular membrane. Cell surface receptors contain an extracellular ligand binding domain, a transmembrane domain and an intracellular domain involved in transmission of signals to the intracellular signal transduction components (Walker, 1994). SERK represents a member of the large group of transmembrane receptor kinases with various functions in plants and animals. Many of these gene products are known to be involved in cellular differentiation processes like Clavata 1 (Clark et al. 1997) or Erecta (Torii et al. 1996). Overexpression or mutation of these genes in plants result in morphological changes in plant organs or plant cells.

The Somatic Embryogenesis Receptor-like Kinase SERK was originally identified as a marker for embryogenic cells, both in vivo, and in vitro. (Schmidt et al. 1997a). Expression of the SERK gene was correlated with the ability to form somatic embryos, a process in which plants are formed from somatic cells through the same morphological, cytological and molecular sequence of stages of embryogenesis as zygotic embryos.

Transmembrane proteins like receptor kinases provide a set of candidate key regulator gene products that are involved in organ or cellular differentiation. In a search for gene products with the ability to modulate the differentiated we searched for receptor-kinase genes expressed in a plant tissues with a large variety of cellular differentiation processes, the influorescense meristem. In a screen for gene products involved in the regulation of the differentiation stage of cells we identified a complete family of receptor-like kinases.

Identification of a new family of receptor-like kinases in Arabidopsis thaliana, the RKS gene family.

In genomic databases of Arabidopsis (accession http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb), a small number of sequences was identified with homology to the Arabidopsis SERK sequence (Schmidt et al. 1997b). These sequences showed homology on nucleotide and predicted amino acid level and were further defined as Receptor Kinases-like SERK (RKS) genes. The initially identified sequences are further defined as RKS_{1-5}. Based on these

five RKS sequences a set of degenerated DNA primers was designed that allowed amplification of possible RKS gene fragments from Arabidopsis.

Primer RKS B forward:

5'-CC[C/G] AAG AT[C/T] AT[A/T] CAC CG[A/C/T] GAT GT[A/C/G] AA[A/G] GC-3'

Primer RKS E reverse

5'-CC[A/G] [A/T]A[A/C/G/T] CC[A/G] AA[A/G] ACA TCG GTT TTC TC-3'

These sequences are based on conserved parts within the nucleotides encoding one exon of the kinase domain. PCR amplification reactions (60 sec. 94°C; 60 sec. 50°C; 90 sec. 72°C) x 40 cycli. were performed with 100 ng of genomic DNA as a template. The resulting PCR products consisted of 209 bp DNA fragments. After cloning in a pGEM-T (Promega) vector, a total of 21 different clones was analysed in order to identify the amplified nucleotide sequences. Removal of the degenerated primer sequences resulted in sequences of 154 nucleotides. Apart from the sequences of RKS1-4 and the SERK gene, a total of 4 new unidentified RKS homologous sequences were identified, further defined as RKS6-10. Sequences from the RKS5 gene were not identified in this screen.

Number of clones isolated and sequenced for different RKS genes followed by time(s) identified in genomic PCR.

RKS1	1
RKS2	4
RKS3	2
RKS4	5
RKS5	0
RKS6	2
RKS7	1
RKS8	2
RKS103	
SERK/RKS0	1

These results indicated the presence of at least 9 different sequences with homology to the conserved kinase domain of the predicted RKS genes (apart from SERK) on the Arabidopsis genome (Figure 1). In order to confirm these data, the fragment of one of the isolated RKS genes was used as a probe in a Southern blot (Figure 2). Low stringency hybridization confirmed the presence of a number of sequences related to the probe fragment. Under the stringency used (see Materials and Methods) a total of approximately 5 hybridizing bands could be observed, indicating the presence of a small RKS gene family in Arabidopsis.

RKS gene expression in Arabidopsis inflorescence tissues.

In order to test whether RKS genes are expressed in tissues where formation of primordia and organs is initiated, RT-PCR reactions were performed on inflorescences. The same combination of PCR primers for RKS fragment amplification was used as described for the genomic PCR reactions. Due to the absence of intron sequences in the described nucleotide fragments, the resulting product was again 209 bp. Starting from the first strand cDNA, a standard PCR reaction was performed for (60 sec. 94°C; 60 sec. 50°C; 90 sec. 72°C) x 40 cycli. In order to obtain a sufficient large amounts of amplified product, a reamplification was performed under similar conditions, using 10% of the mix from the first RT-PCR amplification reactionmix as a template. After cloning in a pGEM-T vector, a total of 21 different clones was sequenced in order to identify the amplified sequences. Removal of the degenerated primer sequences resulted in sequences of 154 nucleotides (Figure 1).

Number of RT-PCR clones isolated and sequenced for different RKS genes followed by time(s) RT-PCR product identified from influorescence tissue:

RKS1	0
RKS2	0
RKS3	2
RKS4	5
RKS5	0
RKS6	0

RKS7	1
RKS8	2
RKS104	
RKS112	
RKS123	
RKS131	
RKS141	
SERK/RKS0	0
RKS	14

These results indicated the presence of at least 14 different sequences with homology to the conserved kinase domain of the predicted RKS genes (apart from SERK) on the Arabidopsis genome (Figure 1). Within influorescenses, at least 9 RKS-like genes were expressed. Within this experiment, expression of RKS 0, 1,2,5 and 6 in inflorescences could not be confirmed. Homology between the different RKS sequences was performed using ALLIGMENT software from Geneworks 2.2 (Figure 3). At least three different subgroups could be visualized of the RKS gene family, representing RKS 2 and RKS6 in subgroup 1, RKS 4, 11, 1, 5,14 and 7 in subgroup 2 and RKS 0, 8, 10, 12 and 13 in subgroup 3. These results confirmed the hybridization patterns, observed with genomic Southerns hybridized with a member of the RKS subgroup 3 (Figure 2). A total of 5 hybridizing bands could be observed, that were likely to represent the genes from RKS 0, 8, 10, 12 and 13.

In order to investigate whether the isolated PCR fragments represented parts of complete RKS genes, full length and partial cDNA clones homologous to these PCR fragments were isolated and characterized.

Isolation and characterization of the RKS gene products in Arabidopsis

A cDNA library from Arabidopsis thaliana Colombia wild type was used to isolate cDNA clones hybridizing with the PCR amplified RKS gene fragments. The consisted of a BRL λZipLox vector containing SalI, NotI linked cDNA inserts from different plant organs (including siliques, flowers, stems, rosette leaves and roots.

Filter hybridization, purification of plaques hybridizing under stringent conditions (65°C, 0.1SSC) with the different RKS fragment probes and finally nucleotide sequence analysis resulted in the characterization of a number of RKS cDNA clones. The predicted amino acid sequences of these clones confirmed that the gene products represent members of the RKS plant receptor kinase family RKS. The sequences from the clones identified by the cDNA library were compared and combined with sequence information from the database http://arabidopsis.org/blast/. Apart from 14 different full length cDNA clones a number of 4 different partial clones were identified.

Overexpression of RKS gene products in transgenic Arabidopsis

Transformation of plasmid DNA into plant cells was performed using A.tumefaciens C58C1. The binary vector used consisted of pGREEN, pGREEN1K or RKS expression constructs. Bacterial colonies were grown on LB agar plates containing 20 mg/L gentamycin, 50 mg/L kanamycin and 50 mg/L rifampicin. Five colonies were used to inoculate 50 ml of LB medium containing 50 mg/L kanamycin and 50 mg/L rifampicin. After 16 hours of incubation at 30°C cells were concentratied by centrifugation and resuspended in 10 ml infiltration medium (consisting of 5% sucrose and 0.05% Silwett L-77 in water. A helper plasmid, necessary for transformation, consisted of the vector pJIC Sa-Rep and was co-transformed together with the pGREEN vector. After electroporation and incubation for 2 hours at 30°C, cells were plated onto LB plates with 50 mg/L rifampicin en 50 mg/L kanamycin. Arabidopsis thaliana wild-type WS cultivar was transformed following the floral dip protocol (Clough and Bent, 1998). In short, the influorescences of young Arabidopsis WS plants grown under long day conditions (16 hours light, 8 hours dark) were dipped for 10 seconds in 10 ml of infiltration solution. Plants were grown further under long day conditions and seeds were harvested after an additional 3-5 weeks. Seeds were surface sterilized in 4% bleach solution for 15 minutes and after extensive washing in sterile water, plated on ½MS plates with 60 mg/L kanamycin. After 10 days incubation under long day conditions, transgenic kanamycin resistent seedlings were isolated and planted on soil for further non-sterile growth under standard

long day greenhouse conditions. This infiltration protocol routinely resulted in approximately 1% transformed seeds for each of the RKS gene constructs used.

Regeneration of Arabidopsis plants after RKS gene transformation

Arabidopsis T2 seeds, obtained from plants infiltrated with A.tumefaciens containing empty pGREEN vectors or pGREEN1K vectors including RKS genes under the control of a 35S promoter, were surface sterilized and added to 40 ml ½MS medium culture to which 1 mg/L 2,4-D was added. After three days of stratification at 4°C, the cultures were incubated on a shaker under long day conditions in a climate room of 20°C for 0-18 days to induce cell proliferation. At different time intervals, seedlings were isolated from the culture, washed and transferred onto ½MS agarplates without 2,4-D or any other hormones. Incubation in the climate room was continued under long day conditions for 4 more weeks. In the absense of RKS genes in the transformed binairy vector, no regeneration of plantlets could be observed (Figure 5C). However, in the presence of RKS gene expression, regenerating plants could be observed that originated from the proliferating cell mass (Figure 5A,B). Different RKS gene constructs showed the ability to regenerate shoot meristems and leaves. The ability to induce regeneration varied between individual integration events and between RKS gene constructs (Figure 5A versus 5B). At this timepoint of 4 weeks of regeneration, plantlets were transferred directly to non-sterile soil and grown for another 4-6 weeks under long day conditions. Fertile, seed setting plants could be obtained from the regenerated plantlets as shown in Figure 5A,B.

20 µg of vector DNA for biolistic DNA delivery into Arabidopsis tissue was mixed with a ballistic suspension mix: 10 mg of gold (Aldrich Chem, Co. Gold 1.5-3 micron), 30 µl 5M NaCl, 5 µl 2M Tris pH 8, 965 µl water, 100 µl 0.1M spermidine, 100 µl 25% PEG, 100 µl 2.5M CaCl2. The suspension was incubated at room temp for 10 min, and centrifuged. The resulting pellet was washed twice with ethanol and resuspended into 200 µl icecold 99.8% ethanol. For each microprojectile bombardment, 10 µl of the gold-coated DNA was used. Bombardment conditions for the HELIUM GUN 461 were: helium pressure 6

bar, vacuum to 50 mbar and 9 cm distance of the tissue from the filter. 0.1 mm mesh size screen was used between tissue and filter, 3 cm distance of the screen from the filter. After bombardment, the Arabidopsis plants were cultured for a period of 3 weeks under long day conditions.

Regeneration in Nicotiana tabacum induced by expression of regeneration-stimulating gene products

20 microgram of plasmid DNA was transferred into cells of tobacco (NTSR1) leaves, using biolistic bombardment with gold particles coated with DNA. Leaf discs were subsequently submerged in liquid MS30 medium (MS medium 30 g sucrose/l, Murashige and Skoog 1962) containing 1 mg/l kinetin and incubated on a rotary shaker (250 rpm) for 14 days. Leaves were then transferred to plates with MS30 plates, 0.8% agar. All incubations have been performed at 20°C with 16 hours light, 8 hours dark. Control experiments with empty or control vectors never gave rise to shoot formation. Regenerating plantlets appeared as a result of particle bombardment with regenerating DNA constructs as shown in figure 6A-C. The transient nature of the introduced construct could be confirmed for 9 out of 10 different regenerants obtained from bombarded tissue (Figure 6D).

Induction of cell proliferation in *Arabidopsis thaliana* induced by expression of regenerion inducing gene products

In order to identify the earlier stages of regeneration after particle bombardment the formation of cellular proliferation was studied as a result of the activity of the regenerating gene product. Single regenerating constructs or combinations of such DNA constructs were bombarded onto two weeks old seedlings of *Arabidopsis thaliana* grown on MS agar plates. Between one and three weeks thereafter the formation of multicellular structures arising from the surface of bombarded rosette leaves could be observed (Figure 6E-H). Bombardments with
empty control vectors never gave rise to the formation of these structures. Interestingly, the proliferating cell mass originating from bombardment with a

GT-W-20S construct developed somatic embryos as a clear indication of regeneration by the process of somatic embryogenesis.
Somatic embryogenesis was hereby not depending on a tissue culture state of the originating tissue but could be directly initiated on adult leaves still attached to the parent plant. Combinations of different regenerating contructs coated on the same gold particle before bombardment allowed also the process of cellular proliferation to be initiated (Figure 6G).Multiple loci of proliferated tissue could be observed on individual leaves after the different regenerating constructs (Figure 6H), indicating that the frequency of regeneration was relatively high when using combinations of regenerating constructs in contrast to bombardments with individual regenerants.

MATERIALS AND METHODS

Southern Blotting

10 µg of genomic DNA from Arabidopsis thaliana wildtype was digested with different restriction enzymes. Fragment DNA was size separated on a 0,9% agarosegel. DNA purination was performed in 0.6M NaCl with 0.4M NaOH. Capillairy blotting was performed onto Hybond N+ membranes. Membranes are hybridized overnight at 65°C in C&G hybridization mix (Church and Gilbert, 1985) and subsequently washed at 65°C with 5SSC, 0,1% SDS. For detection of radioactivity, the Phosphorimager 425 (Molecular Dynamics) was used in combination with phosphoscreen exposure casettes and ImageQuaNT sofware.

DNA fragment purification

DE81 paper (Whatmann) was used for isolation of DNA fragments from agarose gels. Paper segments were introduced into the agarosegel just behind the desired DNA fragments (which were visualized under long wave UV with ethidium bromide staining). Electrophoresis was performed for 10 minutes at 10V/cm gel and the DE81 paper to which the DNA was bound was recovered from the gel. Paper fragments were washed extensively in Low Salt Buffer (LSB) and subsequently DNA was removed from the paper in a small volume of High Salt Buffer (HSB).

LSB (Low Salt Buffer):	HSB (High Salt Buffer):
10 mM Tris pH 7,5	10 mM Tris pH 7,5
1 mM EDTA	1 mM EDTA
100 mM LiCl2	1 M LiCl2
	20% Ethanol

Radioactive Probes

Purified DNA fragments were radiolabelled with 32P-dCTP following a random primed labelling:
50 ng of fragment DNA in 27 µl water is denatured for 5 min. at 100°C. On ice, 21 µl of GAT mix was added: 0,67 M Hepes, 0,17 M Tris, 17 mM MgCl2 ,33 mg/ml acetylated BSA, 25 mg/ml random hexamer primers, 33 mM b-mercapto-ethanol, ,5 mM dNTP's (G + A + T) without dCTP. 2 µl dCTP and 2 µl Klenow (1 U/µl) was added, mixed and incubation was performed for 60 min. at 25°C.

Genomic PCR

Genomic DNA was isolated from wild type Arabidopsis thaliana plants using the protocol of Klimyuk et al. (1993). All PCR reactions were performed in a Thermal Cycler from Perkin Elmer.
PCR amplification reactions were performed under standard conditions using the following mix: 100 ng genomic template DNA in 5 µl water, denatured for 5

min. at 100°C. On ice the following components were added: 2 µl primer B (10 µ M) en 2 ml primer E (10 µM), 1 µl dNTP's (10 mM), 5 µl 10x Taq buffer (Boehringer Mannheim), 0,1 ml Taq polymerase, 5 Units/µl (Boehringer Mannheim), 35 µl water. Paraffin oil was added to the surface in a volume of 20 µl and amplification was performed under the following conditions: (60 sec. 94°C, 60 sec. 50°C, 90 sec. 72°C)x40 cycli. PCR products were routinely purified using the High Pure-PCR product purification kit (Boehringer Mannheim). Purified DNA was cloned in a five-fold molar excess in the pGEM-T Easy vector (Promega) following standard protocols and reaction mixes as supplied within the reaction kit.

RT-PCR

Inflorescences from Arabidopsis thaliana was used as source material to isolate total RNA following the protocol of Siebert and Chenchik (1993)
2.5 µg of total RNA in 10 µl of water was linearized by 1 min. incubation at 100° C, follwed by the addition of the following components on ice:

- 2 µl (10 pmol) dT race primer 5' - GAC TCG AGT CGA CAT CGA TTT TTT TTT TTT TT - 3'
- 1 µl dNTP's (10 mM)
- 4 µl 5x RT buffer (Boehringer Mannheim)
- 0,8 µl reverse transcriptase M-MuLV Expand (Boehringer Mannheim)
- 2 µl 100 mM DTT

Incubation was performed for 60 min. at 42°C, diluted with an equal amount of RNAse free water and stored at -20°C. 2 µl of first strand (= 125 ng) was used in PCR reactions, using the RKS degenerated primers B and E. 2 µl primer B (10 µ M) en 2 µl primer E (10 µM), 1 µl dNTP's (10 mM), 5 µl 10x Taq buffer (Boehringer Mannheim), 0,1 ml Taq polymerase, 5 Units/µl (Boehringer Mannheim), 38 µl water.
Paraffin oil was added to the surface in a volume of 20 µl and amplification was performed under the following conditions: (60 sec. 94°C, 60 sec. 50°C, 90 sec. 72° C)x40 cycli. PCR products were routinely purified using the High Pure-PCR

product purification kit from Boehringer Mannheim. Purified DNA was cloned in a five-fold molar excess in the pGEM-T Easy vector (Promega) following standard protocols and reaction mixes as supplied with the reaction kit.

E.coli and A. tumefaciens transformation

Transformation of plasmid DNA into competent bacteria was performed by electroporation (Dower et al., 1988), using a Genepulser (Biorad). Conditions for electroporation were as follows: 1,5 kV, 25 mF and 200W in standard cuvettes. Directly after transformation, cells were incubated for 90 min. at 37 °C in SOC medium (Sambrook et al. 1989). The bacterial suspension was plated on selective agar plates and incubated overnight at 37°C (E.coli) or for two days at 30°C (A tumefaciens) in order to visualize transgenic bacterial colonies.

Nucleotide sequence analysis

Plasmid DNA was isolated from E.coli by standard boiling method protocol (Sambrook et al. 1989) followed by a subsequent purification with the PCR product purification kit from Boehringer Mannheim. Plasmids were sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Core Kit van Perkin Elmer, using standard protocols as designed for the 480 DNA Thermal Cycler. After electrophoresis on polyacrylamide gels, the results were analysed using the 373A DNA Sequencer from Applied Biosystems. Data were analysed using the software programs Sequencer 3.0, Geneworks 2.2 and DNA-strider 1.2.

cDNA library screening

Plating of the cλZipLox cDNA library was performed as described by the supplier protocols (GIBCO BRL), and plaque lifting and purification as described by Sambrook et al. (1989). cDNA library screening was performed using 20 duplicate filters, each containing approximately 250.000 individual plaques. Filters were screened with different RKS DNA probes representing 209 bp amplified PCR fragment. Prior to labelling, DNA fragments were isolated from the pGEM-T vector by digestion and purified twice by DE81 purification from

agarose gels. Filters were hybridized under stringent conditions (0.1SSC, 65°C). Plaques that hybridized on both filters were isolated and used for two subsequent rounds of further purification. The resulting cDNA clones were sequenced using the T7 and SP6 primers from the primer binding regions of the multiple cloning sit of the λZipLox vector. Internal oligos were designed to sequence the complete cDNA inserts of the RKS clones. Only one cDNA clone was sequenced completely for each RKS gene product identified. An alternative approach to identify and subsequently isolate cDNA clones from RKS genes was to screen the Arabidopsis genome database for RKS homologous sequences and to amplify cDNA clones by RT-PCR approach as described above using primers specific for these RKS gene products, based on the sequence data obtained from Arabidopsis genomic databases (accession http://genome-www2.stanford.edu/cgi-bin/AtDB/nph-blast2atdb). Purified RT-PCR products were cloned in a five-fold molar excess in the pGEM-T Easy vector (Promega) following standard protocols and reaction mixes as supplied with the reaction kit.

Regenerating gene product expression constructs

The CaMV 35S promoter enhanced by duplication of the –343/-90 bp region (Kay et al, 1987) was isolated from the vector pMON999 together with the NOS terminator by NotI digestion. The resulting construct was cloned into the vector pGreen (Bean et al. 1997) and the resulting binairy vector is further defined as pGreen1K. RKS cDNA clones (Figure 2) were isolated from either the pGEM-T easy vector by EcoRI digestion or from the λZipLox vector by EcoRI/BamHI digestion. The resulting cDNA fragments were cloned into respectively EcoRI digested pGreen 1K or EcoR1/BamH1 digested pGreen 1K. Nucleotide sequence analysis was performed in order to test the integrity and the orientation of the RKS cDNA in the vector pGreen1K. The resulting constructs in which the different $RKS_{0\text{-}14}$ had been ligated in the sense configuration with respect to the 35S promoter are further defined as RKS expression constructs. The other regenerating gene products as previously mentioned have been cloned in a similar fashion into the pGreen expression construct under the control of a 35S promoter

Regeneration induced by transient expression of RKS gene products

Rosette leaves and shoot meristems from 3-weeks old Arabdopsis plants grown under long day conditions were surface sterilized in a 1% bleach solution for 20 min, washed extensively with sterile water and placed on ½ MS plates solidified with 0.8% agar.

Particle Bombardment

20 µg of vector DNA for biolistic DNA delivery into plant tissue was mixed with a ballistic suspension mix: 10 mg of gold (Aldrich Chem, Co. Gold 1.5-3 micron), 30 µl

5M NaCl, 5 µl 2M Tris pH 8.0, 965 µl water, 100 µl 0.1M spermidine, 100 µl 25% PEG, 100 µl 2.5M CaCl2. The suspension was incubated at room temp. for 10 min. and centrifuged. The resulting pellet was washed twice with ethanol and resuspended into 200 µl icecold 99.8% ethanol. For each microprojectile bombardment, 10 µl opf the gold-coated DNA was used. Bombardment conditions for the HELIUM GUN 461 were: helium pressure 6 bar, vacuum to 50 mbar and 9 cm distance of the tissue from the filter. 0.1 mm mesh size screen was used between tissue and filter, 3 cm distance of the screen from the filter.

References

Bean SJ, Gooding PS, Mullineaux PM and Davies DR (1997) Plant Cell Reports 16, 513-519.

Blau HM (1989) Trends in Genetics 5, 268-272.

Church C and Gilbert K (1985) Proc. Natl. Acad. Sci USA 81, 1991-1995.

Clark SE, Williams RW and Meyerowitz (1997) Cell 89, 575-585.

Dower WJ et al. (1988) Nucl. Acid Res. 16, 6127-6145.

Kay et al. (1987) Science 236, 1299-1302.

Klimyuk VI, Carroll BJ, Thomas CM and Jones JDG (1993) Plant J. 3, 493-494.

Sambrook J, Fritsch EF and Maniatis T. (1989) Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory, New York.

Schmidt EDL, Hecht V, van Holst GJ, de Vries SC (1997b) production of apomictic seed. International publication number WO97/43427.

Schmidt EDL, Guzzo F, Toonen M, de Vries SC (1997a) Development 124, 2049-2062.

Siebert PD and Chenchik A (1993) Nucl. Acid Res. 21, 2019-2020.

Siegel BA and Verbeke JA 1989, Science 244, 580-582.

Torii KU, Mitsukawa N, Oosumi T, MatsuuraY, Yokoyama R, Whittier RF and Komeda Y (1996) Plant Cell 8, 735-746.

Walker JC (1994) Plant Molecular Biology 26, 1599-1609.

Murashige T. and Skoog F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15, 473-496

Figure legends

Figure 1 depicts the different 154 bp PCR fragments as amplified with the degenerated forward and reverse RKS primers B and E, as shown in Material and Methods. The sequence of the RKS0 fragment is identical with the corresponding region of the Arabidopsis SERK gene. The nucleotide sequences representing the primer sequences have been deleted from the original 209 bp PCR products in this figure.

Figure 2.
Genomic Southern blot of Arabidopsis thaliana genomic DNA digested with different restriction enzymes. 10 µg of genomic digested DNA is loaded in each lane. Low stringency hybridization (65°C, 5SSC) is performed with a 209 bp PCR fragment encoding part of the kinase domain of RKS0.

Figure 3.
Homologies between the 154 bp fragments as amplified from Arabidopsis with the degenerated RKS primers B and E, shown in Figure 1. At least three different subgroups can be visualized of the RKS gene family, representing RKS 2 and RKS6 in subgroup 1, RKS 4, 11, 1, 5,14 and 7 in subgroup 2 and RKS 0, 8, 10, 12 and 13 in subgroup 3. Alignments were performed using DNA Strider 1.2 software.

Figure 4A
Arabidopsis thaliana RKS0 cDNA
The start codon has been indicated by bold capitals.

Figure 4B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-0 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 4 evenly spaced leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and is a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 5

Proliferated cell mass of Arabidopsis plants transformed with different overexpressing constructs of RKS genes (A and B) or with a control pGREEN1K vector without RKS genes. After 18 days of proliferation in the presence of 2,4-D, tissues have been grown for 4 weeks in the absence of hormones. Regenerated plantlets and green shoots are clearly visible in transformed tissues A and B, but absent in the control tissues transformed with the empty pGREEN vector (C).

Figure 6A

Ballistic bombardment of Nicotiana tabacum leaf discs with GT-W-20S at day 0 is followed by a two weeks submerged culture in liquid MS medium 1 mg/L kinetin. Subsequently the discs are cultured on MS agar plates without hormones. Control experiments with empty vector never gave rise to proliferation. The formation of regenerating from leaf explants is shown in days after bombardment.

Figure 6B

Ballistic bombardment of Nicotiana tabacum leaf discs with GT-SBP5-16S at day 0 is followed by a two weeks submerged culture in liquid MS medium with 1mg/L kinetin. Subsequently the leaf discs are cultured on MS agar plates without hormones. The formation of regenerating tissues from leaf explants is shown in days after bombardment. Control experiments with empty vectors never gave rise to shoot formation.

Figure 6C

Nicotiana tabacum callus is bombarded with GT-SBP5-16S at day 0. Callus was generated by incubating tobacco leaves for 6 weeks on MS30, 0.8% agar supplemented with 1mg/L 2,4-D auxin. The callus that formed on the leaves with root like characteristics (extending roots or root hairs from calli) was further cultured on MS30, 0.8% agar petri dishes. The incubation are performed at 20°C with 16 hours light, 8 hours dark. Control experiments with empty vectors never gave rise to shoot formation. 40 days after bombardment regenerating plant can be identified on top of the bombarded callus tissue (plant 1 and plant 2).

Figure 6D

In order to examine the presence of the bombarded DNA regeneration constructs in regenerated plant, tissue samples were taken from 10 different regenerates from the experiments described in the legends of Figure 6A-C. Genomic DNA was isolated from all samples, as well as from two control plants. On this DNA a PRC reaction was performed using primers specific for the NptII gene: construct 1 and 3 from experiment I.

Oligo's used for NptII specific amplification:

Forward oligo: 5'-GCCATGGTGAACAAGATGGATGG-3' Reverse oligo: 5'-GGATCCTCAGAAGAACTCGTCAAG-3'. The resulting PCR product was analysed on agarose gel. Lane 1 and 2 represent regenerates from figure 6C; Lane 3-6 represent regenerates from Figure 6A; Lane 7-10 represent regenerates from Figure 6B. These 10 plants from which tissue material was isolated for lane 1-10 are shown below just prior to DNA isolation. Lane 11 represents a positive control plant that is stable transformed with a control vector (pG1K-GEP). Lane 12 represents a negative control, an untransformed wildtype NTSR1 plant. Lane 13 and 14 represent positive control E.coli purified DNA used for PCR analysis

and M represent marker DNA. Results indicate that only the regenerated plant from lane 8 contained a stable integrated NptII sequence, with all controls giving vector DNA bands.

Figure 6E

Arabidopsis thaliana WS seedlings grown for 14 days on MS agar plates have bombarded with DNA coated gold particles at day 0. Plants are further incubated on the plates at 20°C with 16 hours light, 8 hours dark. Gold particles were coated with 18 microgram of the construct GT-RKS13. In the bombardment procedure, a GUS expression vector was co-bombarded in combination with the GT-W-20S construct in a molar ration of 10% (GUS versus GT-RKS13). Prior to photography, GUS staining was performed on the bombarded tissues. Cell proliferation (arrow) is detectable on the surface of rosette leaves. Control experiments performed with empty vectors did never result in proliferating tissues.

Figure 6F

Ballistic bombardment of Arabidopsis thaliana with GT-W-20S constructs results in cell proliferation on top of the rosette leaver (left). Structures with the morphologic characteristics of somatic embryos appear on the callused structures (middle and right, white arrows). In the bombardment procedure, a GUS expression vector was co-bombarded in combination with the GT-W-20S construct in a molar ration of 10% (GUS versus GT-W-20S). The GT-W-20S construct induces cellular proliferation in neighbouring cells and is unable to induce not contain fragments of the introduced regeneration construct or the GUS expression construct. However, after GUS staining, one cell at the basis of the proliferating cell mass is clearly GUS positive (middle and right, black arrow), indicating that this basal cell has been transformed construct results in the formation of a GUS-negative proliferating cell mass on top of a basal GUS-positive cell. Bombardment studies with empty control vectors did never result in cellular proliferation.

Figure 6G

Ballistic bombardment of Arabidopsis thaliana Ws with GT-CUC2-S, GT-KNAT1-S and GT-CYCD3-S. Cell proliferation becomes already clearly

detectable within one week after bombardment (arrow). Control bombardment studies with empty vectors did not result in cellular proliferation.

Figure 6H

Ballistic bombardment of Arabidopsis thaliana Ws with GT-CUC-2S, GT-KNAT2-S and GT-CYCD3-3S. Different regions of cell proliferation within individual rosette leaves become already clearly detectable within one week after bombardment (arrows). Control bombardment studies with empty vectors did not result in cellular proliferation.

Figure 7

The three different RKS subfamilies I-III based on figure 3. The predicted protein products are shown, and alignment is based on predicted domain structures. Conserved cysteine residues in disulphate bridge formation are underlined.

From the N-terminus towards the C-terminus these domains can be defined as the signal sequence, the extracellular region consisting of respectively a leucine zipper domain, a disulphate bridge domain, an leucine rich repeat domain with 3-5 leucine rich repeats, a putative hydroxyproline domain involved in O-glycosylation, a single transmembrane domain, an intracellular region consisting of respectively an anchor domain, a serine/threonine kinase domain, a domain with unknown function and at the C-terminus a sequence resembling an intracellular leucine rich repeat.

Figure 8A

Arabidopsis thaliana RKS1 cDNA

The start codon has been indicated by bold capitals.

Figure 8B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-1 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids.The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 3 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 9A

Arabidopsis thaliana RKS2 cDNA. The start codon has been indicated by bold capitals.

Figure 9B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-14 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-

proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 10A

Arabidopsis thaliana RKS3 cDNA. The start codon has been indicated by bold capitals.

Figure 10B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-3 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 3 leucine evenly residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation. The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 11A

Arabidopsis thaliana RKS4 cDNA

The start codon has been indicated by bold capitals.

Figure 11B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-4 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 12A

Arabidopsis thaliana RKS5 cDNA. The start codon has been indicated by bold capitals.

Figure 12B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-5 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al.

(1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain has no clear function. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 13A

Arabidopsis thaliana RKS6 cDNA. The start codon has been indicated by bold capitals.

Figure 13B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-6 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.

The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 14A
Arabidopsis thaliana RKS8 cDNA.
The start codon has been indicated by bold capitals.

Figure 14B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-8 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 4 leucine evenly spaced residues, each seperated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 15A

Arabidopsis thaliana RKS10 cDNA. The start codon has been indicated by bold capitals.

Figure 15B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-10 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 4 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.

The seventh domain has an unknown function.

The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.

The ninth domain has an unknown function.

The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 16A

Arabidopsis thaliana RKS11 cDNA/. The start codon has been indicated by bold capitals.

Figure 16B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-11 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al.

(1997). At the predicted extracellular domain the first domain represents a signal sequence. The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation. The fourth domain contains a leucine rich repeat domain, consisting of 3 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 17A
Arabidopsis thaliana RKS12 cDNA. The start codon has been indicated by bold capitals.

Figure 17B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-12 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains

a single transmembrane domain after which the predicted intracellular domains are positioned.

The seventh domain has an unknown function.

The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.

The ninth domain has an unknown function.

The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 18A

Arabidopsis thaliana RKS13 cDNA. The start codon has been indicated by bold capitals.

Figure 18B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-13 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 4 leucine residues, each separated by 7 other amino acids. The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation. The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function. The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The ninth domain has an unknown function. The last and tenth domain at the C-terminal

end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 19A

Arabidopsis thaliana RKS14 cDNA. The start codon has been indicated by bold capitals.

Figure 19B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-14 protein.

Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997). At the predicted extracellular domain the first domain represents a signal sequence. The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.

The third domain contains conserved cysteine residues, involved in disulphate bridge formation. The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues. The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.

The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned. The seventh domain has an unknown function.

The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.

The ninth domain has an unknown function. The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 20 A

Arabidopsis thaliana RKS 7 partial cDNA sequence.

The 5'-end and a region between the two cDNA fragments (.....) is not shown.

Figure 20B

Predicted partial amino acid sequences of the Arabidopsis thaliana RKS-7 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 21 A

Arabidopsis thaliana RKS 9 partial cDNA sequence.

The 5'-end is not shown.

Figure 21B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-9 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 22A

Arabidopsis thaliana RKS 15 partial cDNA sequence.

The 5'-end is not shown.

Figure 22B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-15 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al.

(1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

Figure 23A

Arabidopsis thaliana RKS 16 partial cDNA sequence.
The 5'-end is not shown.

Figure 23B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-16 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

CLAIMS

1. A method for propagation of a plant from plant starting material wherein root and/or shoot initiation is stimulated by introducing at least one recombinant gene product or functional fragment thereof into said starting material allowing reducing or omitting phytohormone addition to said culture.

2. A method according to claim 1 wherein said at least one recombinant gene product or functional fragment thereof is only transiently present in said starting material.

3. A method according to claim 1 or 2 wherein said gene product is derived from a gene involved in the regulation of plant development.

4. A method according to anyone of claims 1 to 3 further comprising transforming at least part of said starting material with a nucleic acid encoding said gene product.

5. A method according to claim 4 wherein said nucleic acid is transiently expressed in said part.

6. A method according to anyone of claims 1 to 5 wherein said culture comprises in vitro culture.

7. A method according to anyone of claims 1 to 6 wherein said propagation comprises essentially seedless propagation.

8. A method according to anyone of claims 1 to 7 wherein said starting material comprises an individual plant cell or protoplast or explant or plant tissue.

9. A method according to anyone of claims 1 to 8 wherein said starting material additionally comprises a recombinant nucleic acid encoding a desired trait.

10. A method according to claim 9 wherein said recombinant nucleic acid encoding a desired trait has additionally been provided with means for nuclear targeting and/or integration in a plant genome.

11. A method according to claim 9 or 10 allowing reducing or omitting selective agent addition to said culture.

12. A method according to anyone of claims 9 to 11 wherein said starting material is devoid of a selectable marker gene conferring resistance to a selective agent.

13. A method according to claim 11 or 12 wherein said selective agent is an antibiotic or an herbicide.

14. A method according to anyone of claims 3 to 13 wherein said gene involved in the regulation of plant development encodes a leucine-rich repeat containing receptor-like kinase.

15. A method according to claim 14 wherein said receptor-like kinase is a representative of a plant receptor kinase family RKS as shown in figure 3.

16. A method according to claim 14 or 15 wherein said receptor-like kinase comprises an N-terminal signal sequence, an extracellular region comprising a leucine zipper domain, a disulphate bridge domain, a leucine rich repeat domain, a proline rich domain, a transmembrane domain, an intracellular region comprising an anchor domain, a serine/trheonine kinase domain and/or a C-terminal leucine rich repeat domain.

17. A method according to anyone of claims 14 to 16 wherein said receptor-like kinase is encoded by a nucleic acid which in *Arabidopsis thaliana* comprises a sequence as shown in anyone of figures 4, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23.

18. A plant or plant material obtainable by a method according to anyone of claims 1 to 17.

19. An isolated and/or recombinant nucleic acid encoding a receptor-like kinase or a functional fragment or functional equivalent thereof, capable of hybridising to a nucleic acid molecule as shown in anyone of figures 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23 or its complementary nucleic acid.

20. A nucleic acid according to claim 19 being at least 75% homologous to a nucleic acid molecule or to a functional equivalent or functional fragment thereof, as shown in anyone of figures 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23, or its complementary nucleic acid.

21. A nucleic acid according to claim 19 or 20 derived from *Arabidopsis thaliana*.

22. A vector comprising a nucleic acid according to anyone of claims 19 to 21.

23. A host cell comprising a nucleic acid according to anyone of claims 19 to 21 or a vector according to claim 22.

24. A nucleic acid according to anyone of claims 19 to 21, a vector according to claim 22 or a host cell according to claim 23 for use in a method according to anyone of claims 1 to 17.

25. An isolated or recombinant proteinaceous substance comprising an amino acid sequence as shown in anyone of figures 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23, or a functional equivalent or functional fragment thereof.

26. A proteinaceous substance according to claim 25 encoded by a nucleic acid according to anyone of claims 19 to 21 or produced by a host cell according to claim 23.

27. A proteinaceous substance according to claim 25 or 26 for use in a method according to anyone of claims 1 to 17.

28. An isolated or synthetic antibody specifically recognising a proteinaceous substance according to claim 25 or 26.

29. An antibody according to claim 28 for use in a method according to anyone of claims 1 to 17.

30. Use of a nucleic acid according to anyone of claims 19 to 21, a vector according to claim 22, a host cell according to claim 23, a proteinaceous substance according to claim 25 or 26 or an antibody according to claim 28 in a method according to anyone of claims 1 to 17.

31. A method for determining a developmental stage of a plant comprising detecting in said plant or parts thereof a nucleic acid according to anyone of claims 19 to 21, or a proteinaceous substance according to claim 25 or 26.

Figure 1 depicts the different 154 bp PCR fragments as amplified with the degenerated forward and reverse RKS primers B and E, as shown in Material and Methods. The sequence of the RKS0 fragment is identical with the corresponding region of the Arabidopsis RKS-0 gene. The nucleotide sequences representing the primer sequences have been deleted from the original 209 bp PCR products in this figure.

RKS1
TCAAGACTGACCCGTGGATAAGTACTCAGGTGCAATGTGGCCAACAGTTCCACGGACTGCAGTTGTGACATGAGAG
TCTTTATGGTCTAGAAGCTTAGCTAACCCGAAATCACCAACAACTGCTTCGAAGTCCTCATCTAACAGAATGTTAG
CT

RKS2
TCAATCATTTCCCTGTGGATATACATTCTGGTGCAATATGACCCATTGTTCCTCGGACCTGAGTGGTTACATTAGTC
TTTTTAACATCTACCAACTTGGCTAAACCAAAATCACCAACCACTGCTTCAAAGTCTTCATCTAGTAACACATTTG
CA

RKS3
ACATCATTTTCCTGTGCAGAGATACTCTGGCGCAATGTGACCCATTGTGCCTCGGACTTGAGTTGTGACATGAGTC
AAACATCTGTCCACAAGCTTAGCTAAACCGAAATCTCCAAGAACTGGCTCAAAATTGTTGTCTAAAAGTATGTTTG
CA

RKS4
ACATCACTGACCAGTGGAGAGATACTCGGGTGCAATGTGACCAACAGTTCCTCTAACCGCGGTTGTGACATGTGAA
TCTTTCTGGTTGAGTAGCTTTGCTAGTCCAAAATCCCCAACAACTGCTTCAAAATACTCATCTAGGAGAATGTTTG
CT

RKS5
TCAAGACTGTCCAGTGGAAAGGTACTCGGGAGCGATGTGTCCAATGGTTCCTCGGACTGCGGTAGTGACATGTGAA
TCTTTCTGGTCTAAAAGCTTTGCTAGACCAAAATCGCCAACTATTGCTTCAAAGCTCTCATCAAGTAGAATATTTG
CA

RKS6
TCATCATTTCCCTGTTGATAAATATTCTGGTGCAATGTGACCCATTGTTCCTCGAACTTGAGTAGTCACATTAGTC
TTTTTAACATCTACTAGCTTGGCTAAACCAAAATCACCAACCACTGCTTCAAAATCTTCATCTAGTAACACGTTAG
CT

RKS7
ACAAGATTGACCAGTTGAGAGATACTCTGGAGCAATGTGACCCACCGTGCCTCTAACCGCGGTTGTCACATGAGAA
TCTTTCATGATCCAAGAGTTTAGCTAAACCAAAATCGCCAACCACAGCTTCACAGTAGTCATCAAGAAGTATATTCG
CT

RKS8
TCAAGATTTTCCAGTTGAGAGATACTCAGGAGCAATGTGTCCAATAGTTCCACGCACAGCCGTTGTGACATGTGTA
TCTTTATAATCCATAAGCCTAGCTAACCCGAAATCACCTACCACCGCCTCAAATTCCTCGTCCAACAGAATATTAG
CA

RKS10
TCATCATTTTCCAGTGGAAAGGTACTCAGGGGCTATATGACCAATTGTCCCACGCACTGCGGTTGTCACATGTGTG
TCTTTTCTAGTCCATGAGTTTTGCAAGTCCAAAATCCCCAACCACGGCTTCAAACTCTTCATCCAACAAAATATTTG
CA

RKS11
ACAAGACTGACCAGTGGAGAGATATTCAGGTGCAATGTGGCCAACCGTACCACGGACCGCAGTTGTGACATGAGAA
TCTTGCATGGTTAAGGAGCTTTGCGAGTCCAAAGTCACCAACAACAGCTTCAAAGCACTCGTCTAAGAGAATATTAG
CT

RKS12
ACAAGATTTTCCTGTCGAGAGGTACTCGGGAGCTATATGGCCAATCGTACCGCGTACAGCAGTTGTCACATGGGAG
TCATTGTAATTCATTAATTTTGCTAGCCCAAAGTCTCCAACAACAGCTTCAAACTCTTCATCTAACAGTATATTTG
CA

RKS13

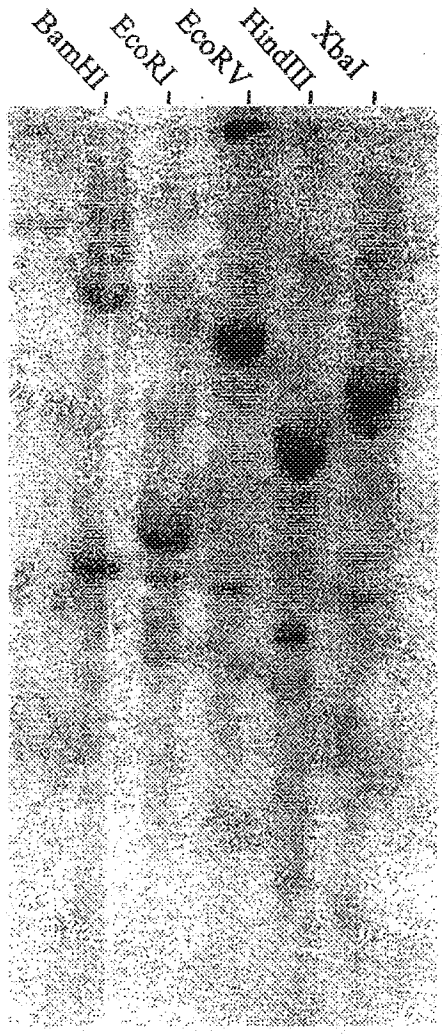
FIGUUR 1 CONTD.

TGCTAATATATTGTTAGATGAAGAGTTTGAAGCTGTTGTTGGAGATTTTGGGCTCGCAAAATTAATGAATTATAAT
GACTCCCATGTGACAACTGCTGTACGCGGTACAATTGGCCATATAGCGCCCGAGTACCTCTCGACAGGAAAATCTT
CT

RKS14
TGCGAACATACTTCTTGACGATTACTTTGAAGCTGTTGTCGGAGATTTCGGGTTGGCTAAGCTTTTGGATCATGAG
GAGTCGCATGTGACAACCGCCGTGAGAGGAACAGTGGGTCACATTGCACCTGAGTATCTCTCAACAGGACAATCTT
CT

RKS0
TGAAGATTTTCCGGTTGAGAGATATTCTGGAGCGATGTGACCGATGGTGCCACGGACTGCTGTTGTCACGTGAGTG
TCTTTATAGTCCATAAGCTTTGCCAACCCGAAATCTCCAACAACCGCTTCGAATTCTTCGTCTAAGAGGATGTTTG
CT

FIGURE 2



5 x SSC

FIGURE 3

ALLIGNMENT UPGMA Tree

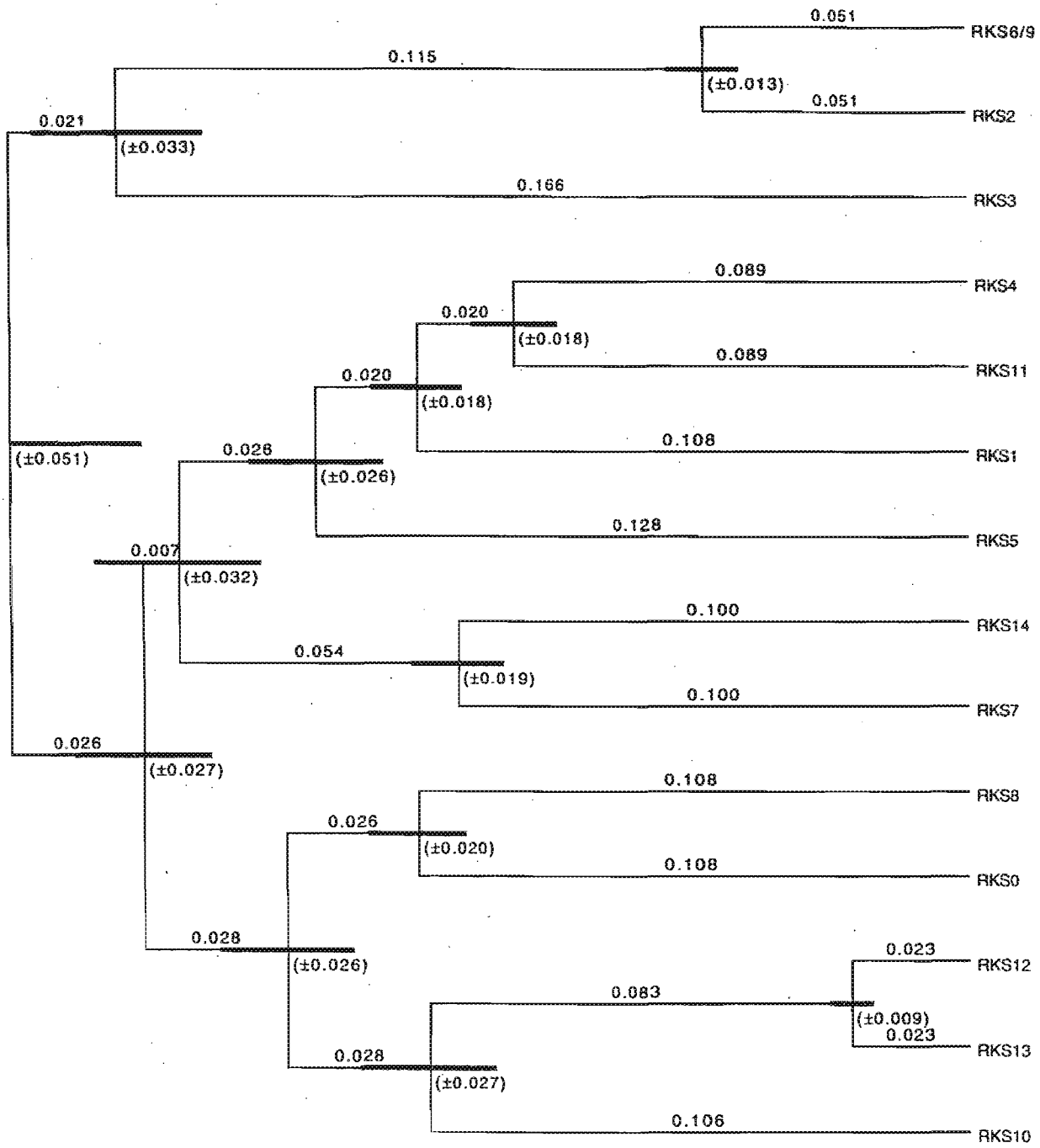


Figure 4a
Arabidopsis thaliana RKS0 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                                    31/11
att ttt att tta ttt ttt act ctt tgt ttg ttt taa tgc taa tgg gtt ttt aaa agg gtt

61/21                                                  91/31
atc gaa aaa atg agt gag ttt gtg ttg agg ttg tct ctg taa agt gtt aat ggt ggt gat

121/41                                                 151/51
ttt cgg aag tta ggg ttt tct cgg atc tga aga gat caa atc aag att cga aat tta cca

181/61                                                 211/71
ttg ttg ttt gaa ATG GAG TCG AGT TAT GTG GTG TTT ATC TTA CTT TCA CTG ATC TTA CTT

241/81                                                 271/91
CCG AAT CAT TCA CTG TGG CTT GCT TCT GCT AAT TTG GAA GGT GAT GCT TTG CAT ACT TTG

301/101                                                331/111
AGG GTT ACT CTA GTT GAT CCA AAC AAT GTC TTG CAG AGC TGG GAT CCT ACG CTA GTG AAT

361/121                                                391/131
CCT TGC ACA TGG TTC CAT GTC ACT TGC AAC AAC GAG AAC AGT GTC ATA AGA GTT GAT TTG

421/141                                                451/151
GGG AAT GCA GAG TTA TCT GGC CAT TTA GTT CCA GAG CTT GGT GTG CTC AAG AAT TTG CAG

481/161                                                511/171
TAT TTG GAG CTT TAC AGT AAC AAC ATA ACT GGC CCG ATT CCT AGT AAT CTT GGA AAT CTG

541/181                                                571/191
ACA AAC TTA GTG AGT TTG GAT CTT TAC TTA AAC AGC TTC TCC GGT CCT ATT CCG GAA TCA

601/201                                                631/211
TTG GGA AAG CTT TCA AAG CTG AGA TTT CTC CGG CTT AAC AAC AAC AGT CTC ACT GGG TCA

661/221                                                691/231
ATT CCT ATG TCA CTG ACC AAT ATT ACT ACC CTT CAA GTG TTA GAT CTA TCA AAT AAC AGA

721/241                                                751/251
CTC TCT GGT TCA GTT CCT GAC AAT GGC TCC TTC TCA CTC TTC ACA CCC ATC AGT TTT GCT

781/261                                                811/271
AAT AAC TTA GAC CTA TGT GGA CCT GTT ACA AGT CAC CCA TGT CCT GGA TCT CCC CCG TTT

841/281                                                871/291
TCT CCT CCA CCA CCT TTT ATT CAA CCT CCC CCA GTT TCC ACC CCG AGT GGG TAT GGT ATA

901/301                                                931/311
ACT GGA GCA ATA GCT GGT GGA GTT GCT GCA GGT GCT GCT TTG CCC TTT GCT GCT CCT GCA

961/321                                                991/331
ATA GCC TTT GCT TGG TGG CGA CGA AGA AGC CCA CTA GAT ATT TTC TTC GAT GTC CCT GCC

1021/341                                               1051/351
GAA GAA GAT CCA GAA GTT CAT CTG GGA CAG CTC AAG AGG TTT TCT TTG CGG GAG CTA CAA

1081/361                                               1111/371
GTG GCG AGT GAT GGG TTT AGT AAC AAG AAC ATT TTG GGC AGA GGT GGG TTT GGG AAA GTC

1141/381                                               1171/391
TAC AAG GGA CGC TTG GCA GAC GGA ACT CTT GTT GCT GTC AAG AGA CTG AAG GAA GAG CGA

1201/401                                               1231/411
ACT CCA GGT GGA GAG CTC CAG TTT CAA ACA GAA GTA GAG ATG ATA AGT ATG GCA GTT CAT

1261/421                                               1291/431
CGA AAC CTG TTG AGA TTA CGA GGT TTC TGT ATG ACA CCG ACC GAG AGA TTG CTT GTG TAT

1321/441                                               1351/451
CCT TAC ATG GCC AAT GGA AGT GTT GCT TCG TGT CTC AGA GAG AGG CCA CCG TCA CAA CCT
```

FIGUUR 4a CONTD.

1381/461 1411/471
CCG CTT GAT TGG CCA ACG CGG AAG AGA ATC GCG CTA GGC TCA GCT CGA GGT TTG TCT TAC

1441/481 1471/491
CTA CAT GAT CAC TGC GAT CCG AAG ATC ATT CAC CGT GAC GTA AAA GCA GCA AAC ATC CTC

1501/501 1531/511
TTA GAC GAA GAA TTC GAA GCG GTT GTT GGA GAT TTC GGG TTG GCA AAG CTT ATG GAC TAT

1561/521 1591/531
AAA GAC ACT CAC GTG ACA ACA GCA GTC CGT GGC ACC ATC GGT CAC ATC GCT CCA GAA TAT

1621/541 1651/551
CTC TCA ACC GGA AAA TCT TCA GAG AAA ACC GAC GTT TTC GGA TAC GGA ATC ATG CTT CTA

1681/561 1711/571
GAA CTA ATC ACA GGA CAA AGA GCT TTC GAT CTC GCT CGG CTA GCT AAC GAC GAC GAC GTC

1741/581 1771/591
ATG TTA CTT GAC TGG GTG AAA GGA TTG TTG AAG GAG AAG AAG CTA GAG ATG TTA GTG GAT

1801/601 1831/611
CCA GAT CTT CAA ACA AAC TAC GAG GAG AGA GAA CTG GAA CAA GTG ATA CAA GTG GCG TTG

1861/621 1891/631
CTA TGC ACG CAA GGA TCA CCA ATG GAA AGA CCA AAG ATG TCT GAA GTT GTA AGG ATG CTG

1921/641 1951/651
GAA GGA GAT GGG CTT GCG GAG AAA TGG GAC GAA TGG CAA AAA GTT GAG ATT TTG AGG GAA

1981/661 2011/671
GAG ATT GAT TTG AGT CCT AAT CCT AAC TCT GAT TGG ATT CTT GAT TCT ACT TAC AAT TTG

2041/681 2071/691
CAC GCC GTT GAG TTA TCT GGT CCA AGG taa aaa aaa aaa aaa aaa aa 2087
Stop

Figure 4B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-0 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 4 evenly spaced leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and is a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MESSYVVFILLSLILLPNHSL
WLASANLEG

DALHTLRVTLVDP
NNVLQSWDPTLVN

PCTWFHVTCNNENSVIRV

DLGNAELSGHLV
P ELGVLKNLQYLELYSNNITGPI
PSNLGNLTNLVSLDLYLNSFSGPI
PESLGKLSKLRFLRLNNNSLTGSI
PMSLTNITTLQVLDLSNNRLSGSV
PDNGSFSLFTPISFANNLDLCGPV

TSHPCPGSPPFSPPPP
FIQPPPVSTPSGYGITG

AIAGGVAAGAAL
PFAAPAIAFAWW

RRRSPLDIFFDVPAEEDPE
VHLGQLKRFSLRELQVAS

DGFSNKNILGRGGFGKVYKGRLAD
GTLVAVKRLKEERTPGGELQFQ
TEVEMISMAVHRNLLRLRGFCM
TPTERLLVYPYMANGSVASCLR
ERPPSQPPLDWPTRKRIALGSA
RGLSYLHDHCDPKIIHRDVKAA
NILLDEEFEAVVGDFGLAKLMD
YKDTHVTTAVRGTIGHIAPEYL
STGKSSEKTDVFGYGIMLLELI
TGQRAFDLARLANDDDVMLLDW
VKGLLKEKKLEMLVDPDLQTNY
EERELEQVIQVALLCTQGSPME
RPKMSEVVRMLE

GDGLAEKWDEWQKVEILREEIDLS

PNPNSDWILDSTYNLHAVELSGPR (625)

FIGURE 5

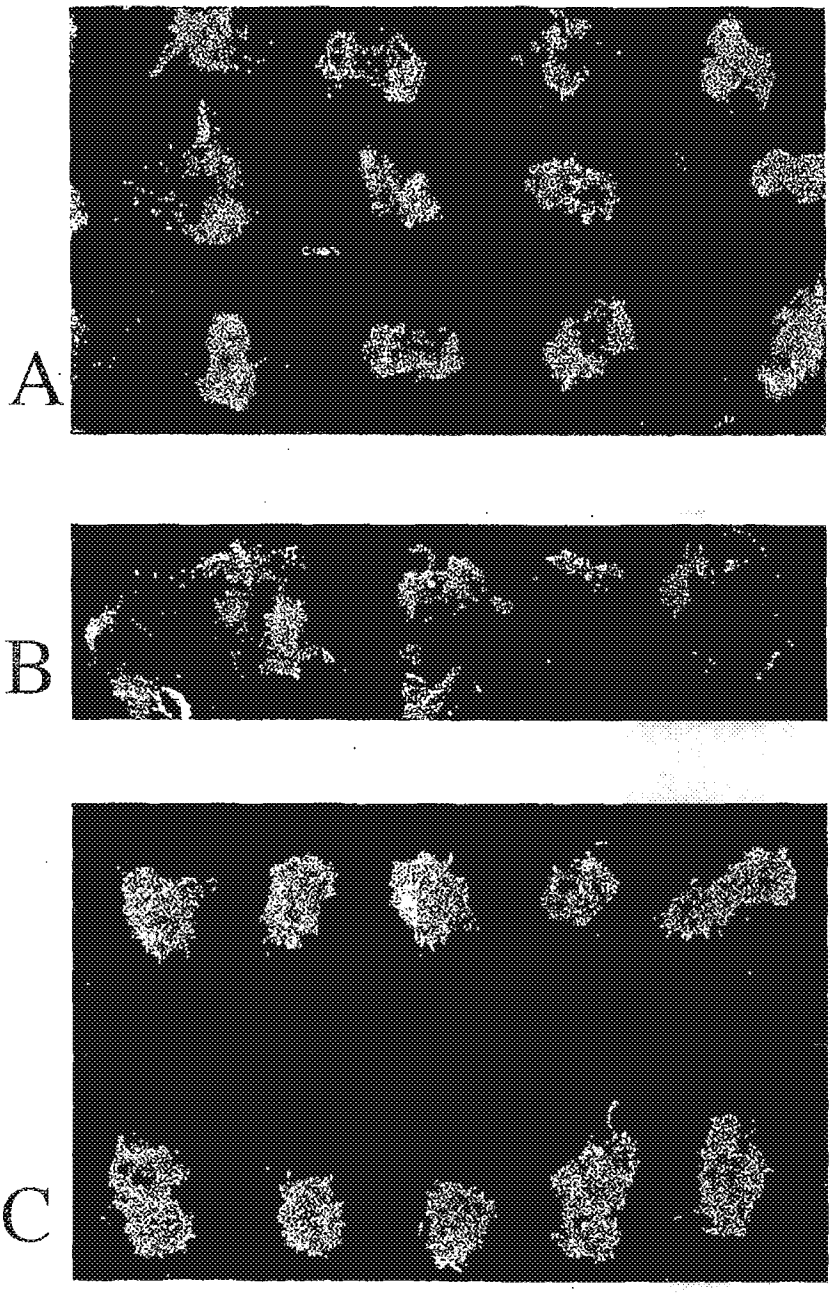
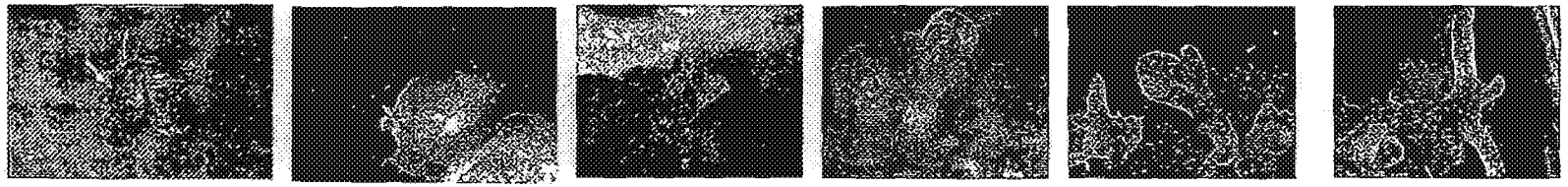
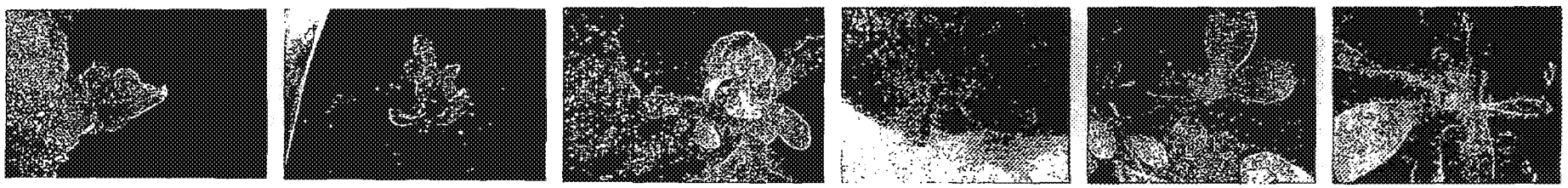


Figure 6A
Ballistic bombardment of Nicotiana tabacum leaf discs with GT-W-20S at day 0 is followed by a two weeks submerged culture in liquid MS medium with 1 mg/L kinetin. Subsequently the leaf discs are cultured on MS agar plates without hormones. Control experiments with empty vector never gave rise to proliferation. The formation of regenerating tissues from leaf explants
is shown in days after bombardment.



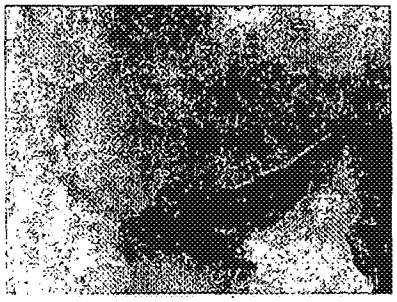
day 44 day 55 day 64 day 69 day 71 day 100

Figure 6B
Ballistic bombardment of Nicotiana tabacum leaf discs with GT-SBP5-16S at day 0 is followed by a two weeks submerged culture in liquid MS medium with 1 mg/L kinetin. Subsequently the leaf discs are cultured on MS agar plates without hormones. The formation of regenerating tissues from leaf explants is shown in days after bombardment. Control experiments with empty vectors never gave rise to shoot formation.

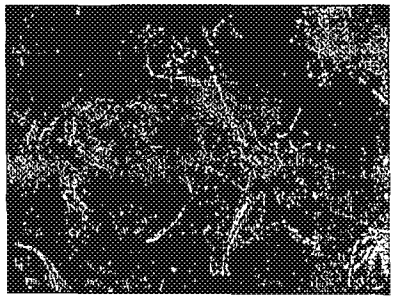


day 42 day 50 day 55 day 64 day 69 day 100

Figure 6C
Nicotiana tabacum callus is bombarded with GT-SBP5-16S at day 0. Callus was generated by incubating tobacco leaves for 6 weeks on MS30, 0.8% agar supplemented with 1 mg/L 2,4-D auxin. The callus that formed on the leaves with root like characteristics (extending roots or root hairs from calli) was further cultured on MS30, 0.8% agar petri dishes. The incubation are performed at 20°C with 16 hours light, 8 hours dark. Control experiments with empty vectors never gave rise to shoot formation. 40 days after bombardment regenerating plants can be identified on top of the bombarded callus tissue (plant 1 and plant 2).



plant 1



plant 2

Figure 6D

In order to examine the presence of the bombarded DNA regeneration constructs in regenerated plants, tissue samples were taken from 10 different regenerants from the experiments described in the legends of Figure 6A-C. Genomic DNA was isolated from all samples, as well as from two control plants.

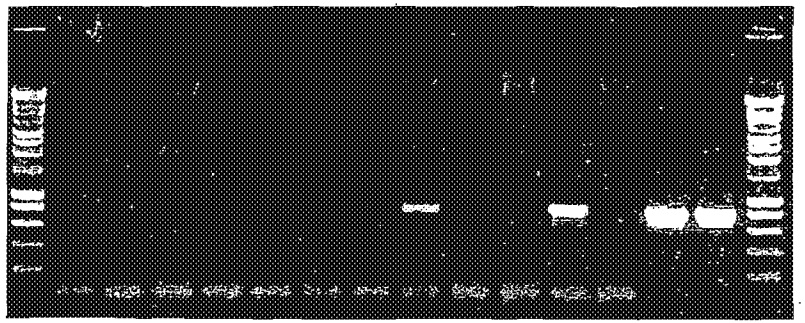
On this DNA a PCR reaction was performed using primers specific for the NptII gene, which was located on the plasmid used for particle bombardment. As a control the PCR was also performed on two plasmid DNA's containing the NptII gene, construct 1 and 3 from experiment I.

Oligo's used for NptII specific amplification:

Forward oligo: 5'-GCCATGGTTGAACAAGATGGATGG-3' Reverse oligo: 5'-GGATCCTCAGAAGAACTCGTCAAG-3'.

The resulting PCR product was analyzed on agarose gel. Lane 1 and 2 represent regenerants from Figure 6C; Lane 3-6 represent regenerants from Figure 6A; Lane 7-10 represent regenerants from Figure 6B. These 10 plants from which tissue material was isolated for lane 1-10 are shown below just prior to DNA isolation. Lane 11 represents a positive control plant that is stable transformed with a control vector (pG1K-GFP). Lane 12 represents a negative control, an untransformed wildtype NTSR1 plant. Lane 13 and 14 represent positive control E.coli purified DNA used for PCR analysis and M represent marker DNA. Results indicate that only the regenerated plant from lane 8 contained a stable integrated NptII sequence, with all controls giving expected vector DNA bands.

plant 1 plant 2 plant 3 plant 4 plant 5 plant 6 plant 7 plant 8 plant 9 plant 10



M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 M

Figure 6E

Arabidopsis thaliana WS seedlings grown for 14 days on MS agar plates have bombarded with DNA coated gold particles at day 0. Plants are further incubated on the plates at 20°C with 16 hours light, 8 hours dark. Gold particles were coated with 18 microgram of the construct GT-RKS13. In the bombardment procedure, a GUS expression vector was co-bombarded in combination with the GT-W-20S construct in a molar ration of 10% (GUS versus GT-RKS13). Prior to photography, GUS staining was performed on the bombarded tissues. Cell proliferation (arrow) is detectable on the surface of rosette leaves. Control experiments performed with empty vecors did never result in proliferating tissues.

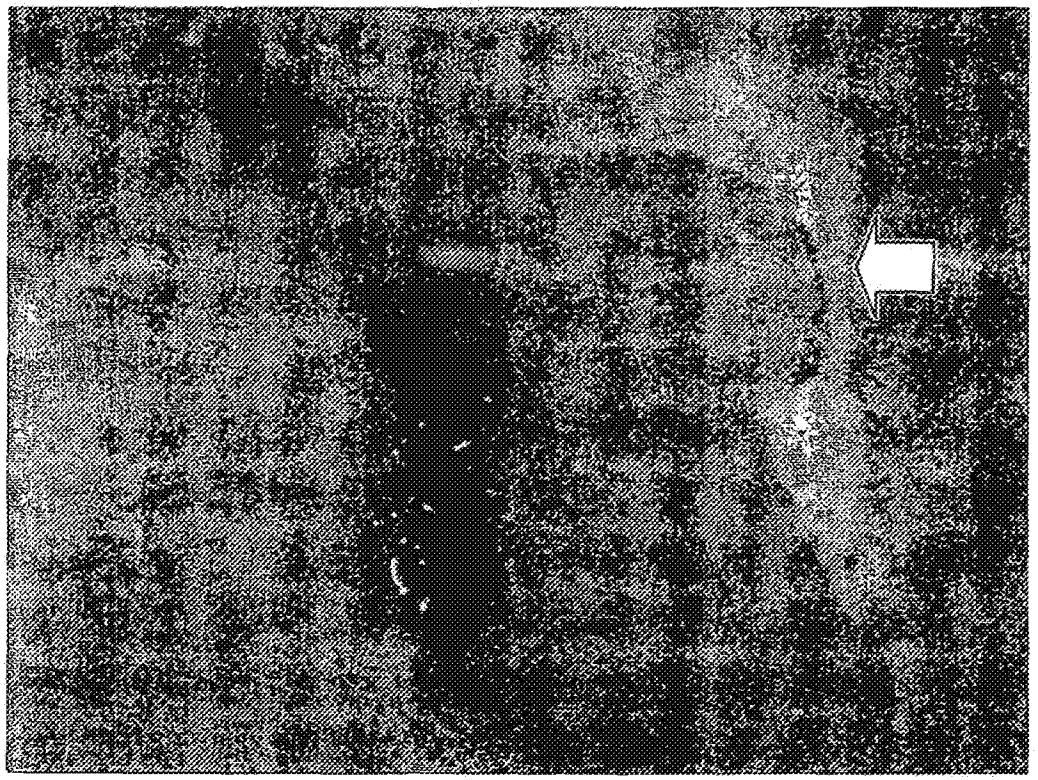


Figure 6F
Ballistic bombardment of *Arabidopsis thaliana* with GT-W-20S constructs results in cell proliferation on top of the rosette leaves (left). Structures with the morphologic characteristics of somatic embryos appear on the surface of the callused structures (middle and right, white arrows). In the bombardment procedure, a GUS expression vector was co-bombarded in combination with the GT-W-20S construct in a molar ration of 10% (GUS versus GT-W-20S). The GT-W-20S construct induces cellular proliferation in neighbouring cells and is unable to induce cellular proliferation of (de)-differentiation of the expressing cell itself. The resulting proliferating cell mass is therefore untransformed and does not contain fragments of the introduced regeneration construct or the GUS expression construct. However, after GUS staining, one cell at the basis of the proliferating cell mass is clearly GUS positive (middle and right, black arrow), indicating that this basal cell has been transformed with the bombarded constructs. A similar process might have occured as shown in figure 6E, where the GT-RKS13 introduced expression construct results in the formation of a GUS-negative proliferating cell mass on top of a basal GUS-positive cell.
Bombardment studies with empty control vectors did never result in cellular proliferation.

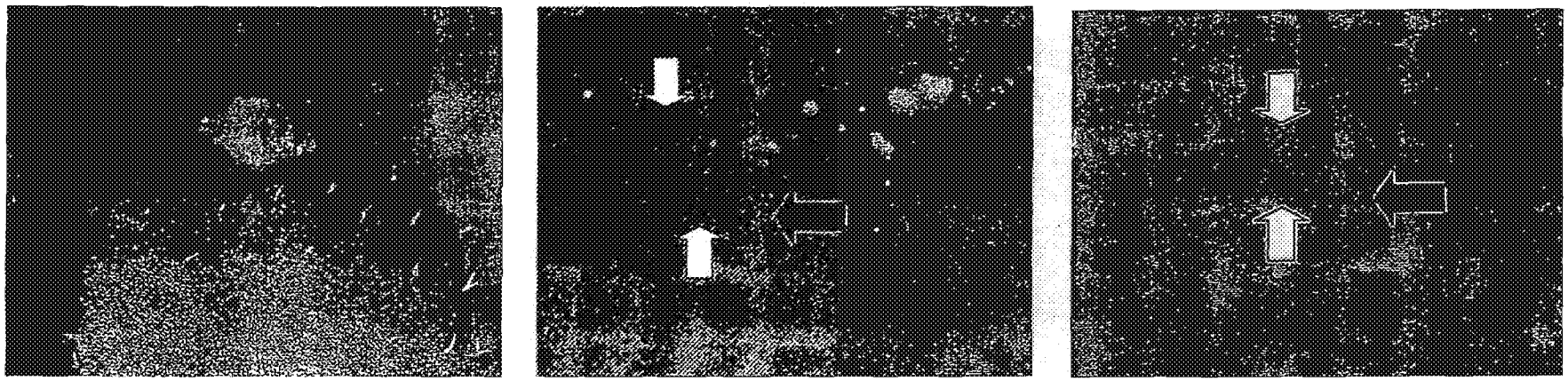


Figure 6G
Ballistic bombardment of *Arabidopsis thaliana* WS with GT-CUC2-S, GT-KNAT1-S and GT-CYCD3-S. Cell proliferation becomes already clearly detectable within one week after bombardment (arrow). Control bombardment studies with empty vectors did not result in cellular proliferation.

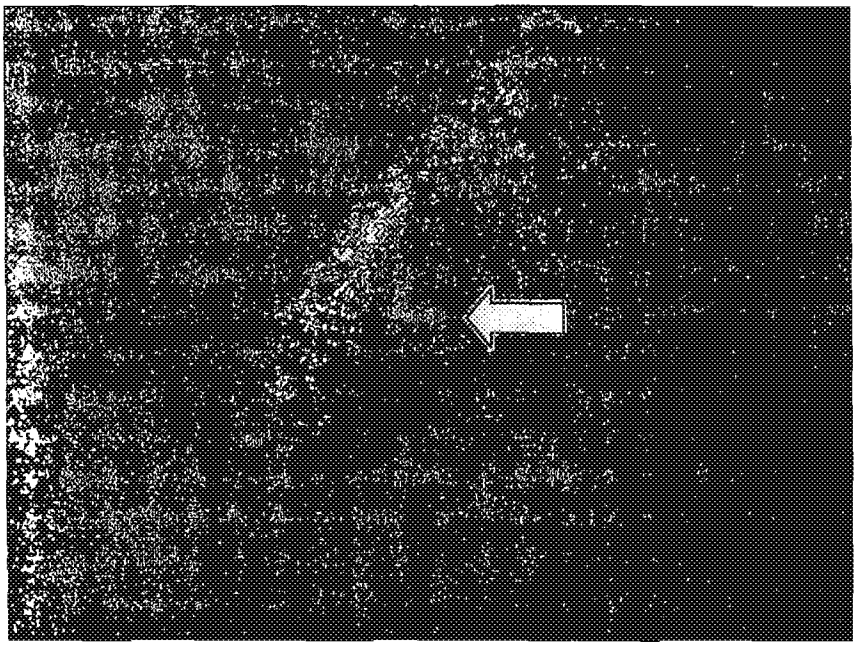


Figure 6H

Ballistic bombardment of *Arabidopsis thaliana* WS with GT-CUC-2S, GT-KNAT2-S and GT-CYCD3-3S. Different regions of cell proliferation within individual rosette leaves become already clearly detectable within one week after bombardment (arrows). Control bombardment studies with empty vectors did not result in cellular proliferation.

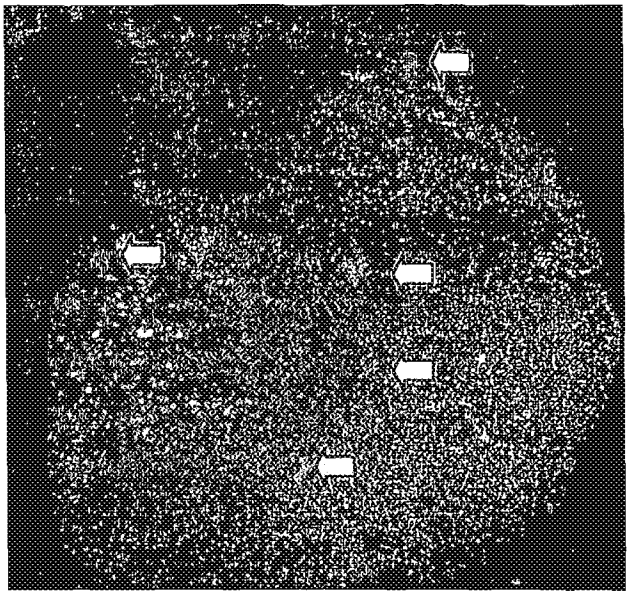


FIGURE 7

Figure 7. Predicted protein domains of the RKS subfamily I

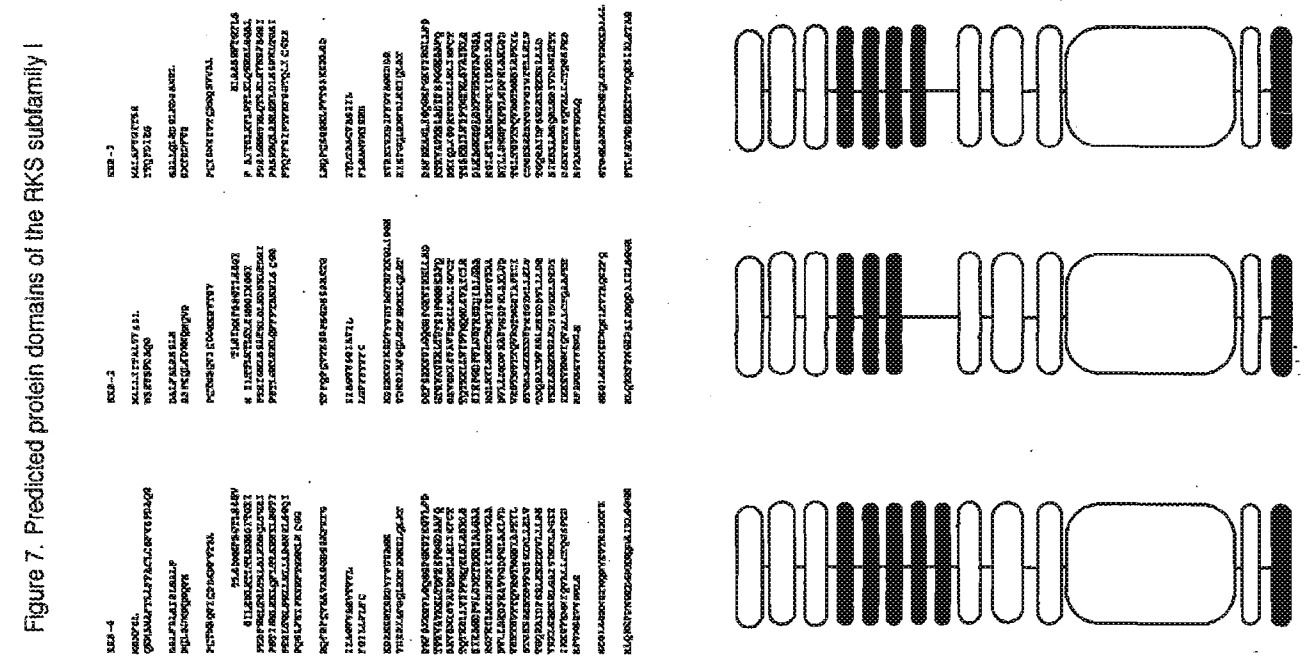


FIGURE 7

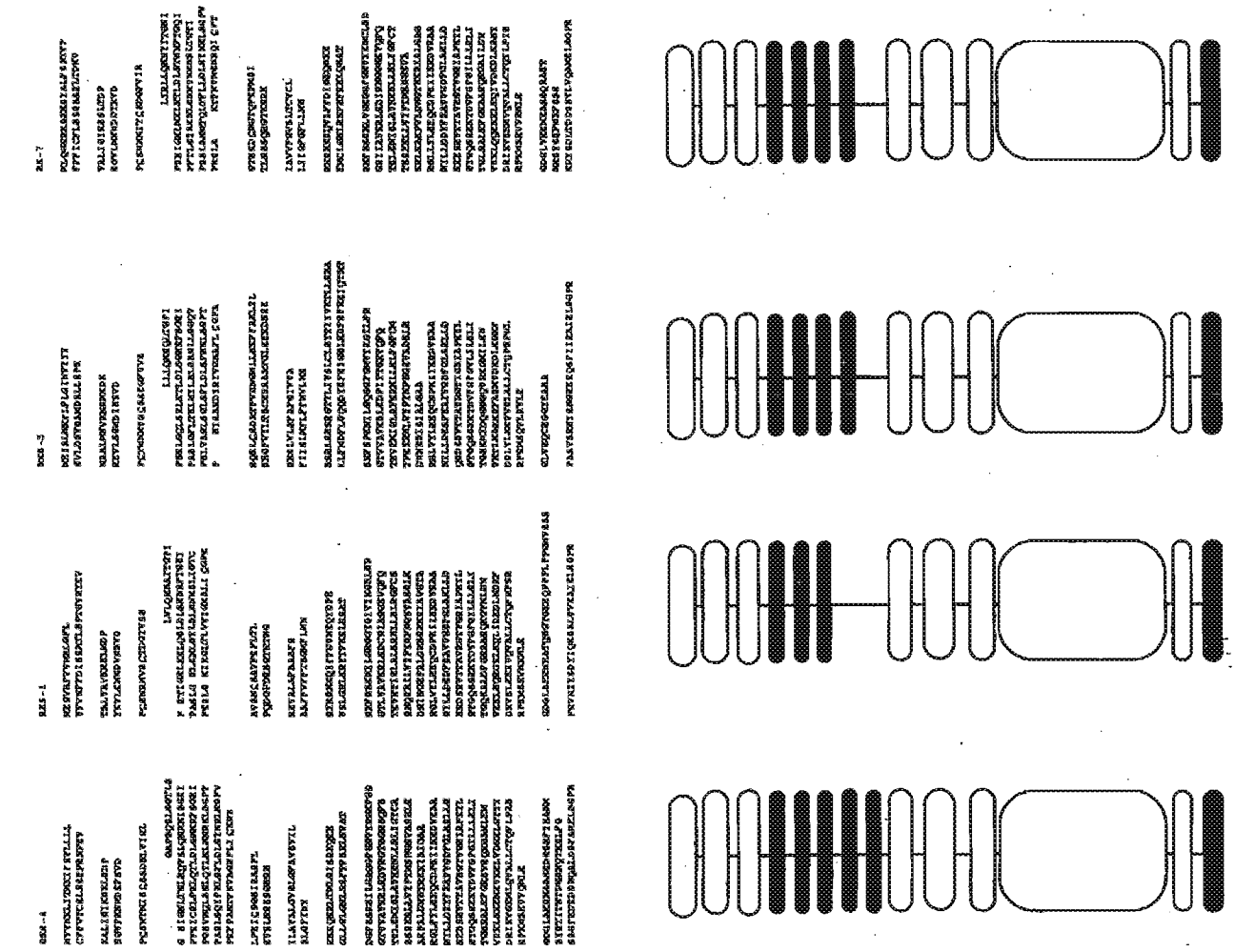


FIGURE 7

Figure 7. Predicted protein domains of the RKS subfamily III

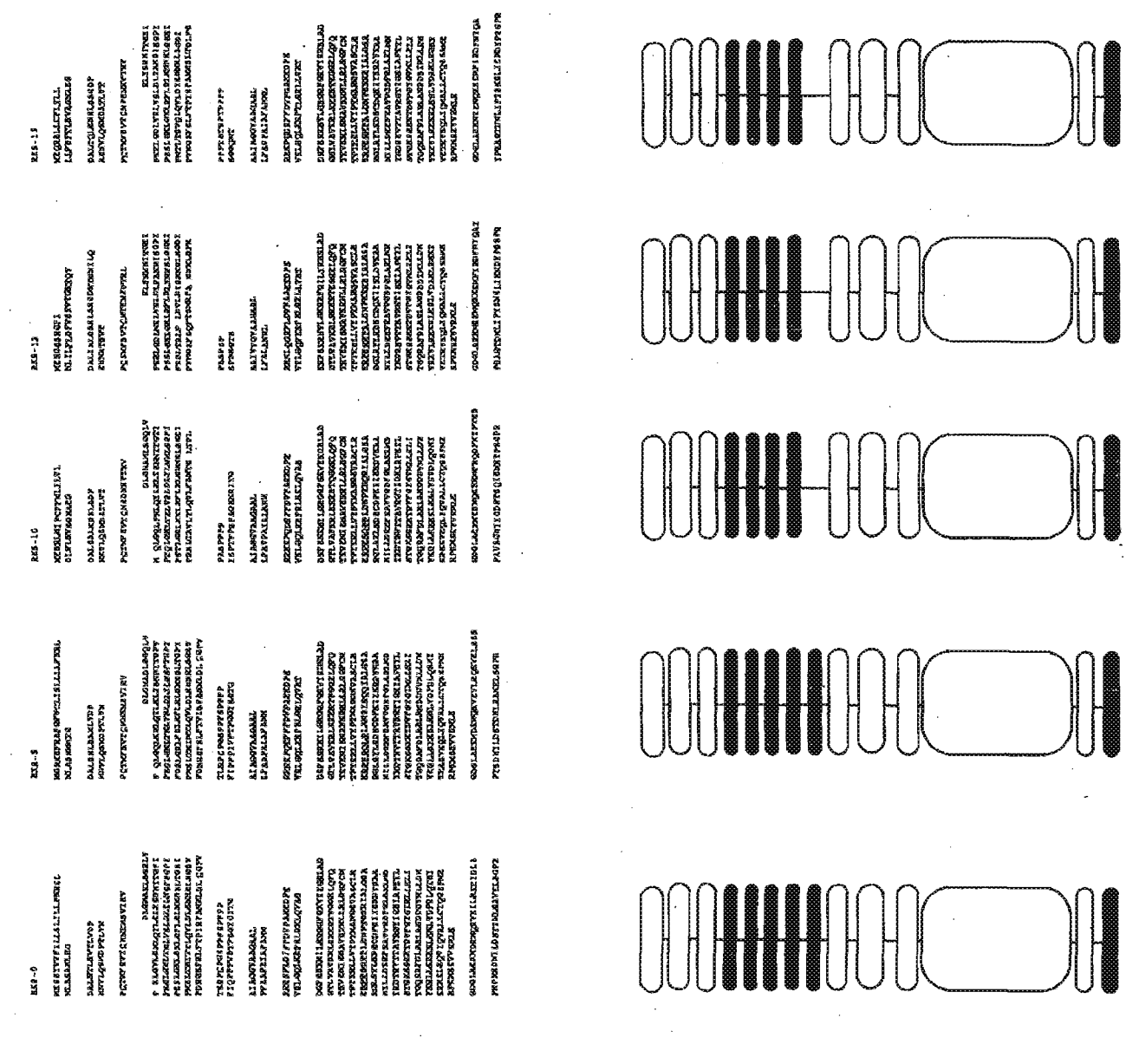


Figure 8a
Arabidopsis thaliana RKS1 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                     31/11
cca aag ttg att gct tta aga agg gat ATG gaa ggt gtg aga ttt gtg gtg tgg aga tta

61/21                                   91/31
gga ttt ctg gtt ttt gta tgg ttc ttt gat atc tct tct gct aca ctt tct cct act ggt

121/41                                  151/51
gta aac tat gaa gtg aca gct ttg gtt gct gtg aag aat gaa ttg aat gat ccg tac aaa

181/61                                  211/71
gtt ctt gag aat tgg gat gtg aat tca gtt gat cct tgt agc tgg aga atg gtt tct tgc

241/81                                  271/91
act gat ggc tat gtc tct tca ctg gtg ttg caa aac aat gca atc act ggt cca att ccg

301/101                                 331/111
gaa acg att ggg agg ttg gag aag ctt cag tca ctt gat ctt tcg aac aat tca ttc acc

361/121                                 391/131
ggg gag ata ccg gcc tca ctt gga gaa ctc aag aac ttg aat tac ttg cgg tta aac aat

421/141                                 451/151
aac agt ctt ata gga act tgc cct gag tct cta tcc aag att gag gga ctc act cta gtg

481/161                                 511/171
gta att ggt aat gcg tta atc tgt ggc cca aaa gct gtt tca aac tgt tct gct gtt ccc

541/181                                 571/191
gag cct ctc acg ctt cca caa gat ggt cca gat gaa tca gga act cgt acc aat ggc cat

601/201                                 631/211
cac gtt gct ctt gca ttt gcc gca agc ttc agt gca gca ttt ttt gtt ttc ttt aca agc

661/221                                 691/231
gga atg ttt ctt tgg tgg aga tat cgc cgt aac aag caa ata ttt ttt gac gtt aat gaa

721/241                                 751/251
caa tat gat cca gaa gtg agt tta ggg cac ttg aag agg tat aca ttc aaa gag ctt aga

781/261                                 811/271
tct gcc acc aat cat ttc aac tcg aag aac att ctc gga aga ggc gga tac ggg att gtg

841/281                                 871/291
tac aaa gga cac tta aac gat gga act ttg gtg gct gtc aaa cgt ctc aag gac tgt aac

901/301                                 931/311
att gcg ggt gga gaa gtc cag ttt cag aca gaa gta gag act ata agt ttg gct ctt cat

961/321                                 991/331
cgc aat ctc ctc cgg ctc cgc ggt ttc tgt agt agc aac cag gag aga att tta gtc tac

1021/341                                1051/351
cct tac atg cca aat ggg agt gtc gca tca cgc tta aaa gat aat atc cgt gga gag cca

1081/361                                1111/371
gca tta gac tgg tcg aga agg aag aag ata gcg gtt ggg aca gcg aga gga cta gtt tac

1141/381                                1171/391
cta cac gag caa tgt gac ccg aag att ata cac cgc gat gtg aaa gca gct aac att ctg

1201/401                                1231/411
tta gat gag gac ttc gaa gca gtt gtt ggt gat ttt ggg tta gct aag ctt cta gac cat

1261/421                                1291/431
aga gac tct cat gtc aca act gca gtc cgt gga act gtt ggc cac att gca cct gag tac

1321/441                                1351/451
tta tcc acg ggt cag tcc tca gag aag act gat gtc ttt ggc ttt ggc ata ctt ctc ctt
```

FIGUUR 8a CONTD.

1381/461 1411/471
gag ctc att act ggt cag aaa gct ctt gat ttt ggc aga tcc gca cac cag aaa ggt gta

1441/481 1471/491
atg ctt gac tgg gtg aag aag ctg cac caa gaa ggg aaa cta aag cag tta ata gac aaa

1501/501 1531/511
gat cta aat gac aag ttc gat aga gta gaa ctc gaa gaa atc gtt caa gtt gcg cta ctc

1561/521 1591/531
tgc act caa ttc aat cca tct cat cga ccg aaa atg tca gaa gtt atg aag atg ctt gaa

1621/541 1651/551
ggt gac ggt ttg gct gag aga tgg gaa gcg acg cag aac ggt act ggt gag cat cag cca

1681/561 1711/571
ccg cca ttg cca ccg ggg atg gtg agt tct tcg ccg cgt gtg agg tat tac tcg gat tat

1741/581 1771/591
att cag gaa tcg tct ctt gta gta gaa gcc att gag ctc tcg ggt cct cga tga

Figure 8b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-1 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 3 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MEGVRFVVWRLGFL
VFVWFFDISSATLSPTGVNYEV

TALVAVKNELNDP
YKVLENWDVNSVD

PCSWRMVSCTDGYVSS

LVLQNNAITGPI
P ETIGRLEKLQSLDLSNNSFTGEI
PASLG ELKNLNYLRLNNNSLIGTC
PESLS KIEGLTLVVIGNALICGPK

AVSNCSAVPEPLTL
PQDGPDESGTRTNG

HHVALAFAASFS
AAFFVFFTSGMFLWW

RYRRNKQIFFDVNEQYDPE
VSLGHLKRYTFKELRSAT

NHFNSKNILGRGGYGIVYKGHLND
GTLVAVKRLKDCNIAGGEVQFQ
TEVETISLALHRNLLRLRGFCS
SNQERILVYPYMPNGSVASRLK
DNIRGEPALDWSRRKKIAVGTA
RGLVYLHEQCDPKIIHRDVKAA
NILLDEDFEAVVGDFGLAKLLD
HRDSHVTTAVRGTVGHIAPEYL
STGQSSEKTDVFGFGILLLELI
TGQKALDFGRSAHQKGVMLDW
VKKLHQEGKLKQLIDKDLNDKF
DRVELEEIVQVALLCTQFNPSH
RPKMSEVMKMLE

GDGLAERWEATQNGTGEHQPPPLPPGMVSSS

PRVRYYSDYIQESSLVVEAIELSGPR

Figure 9a
Arabidopsis thaliana RKS2 cDNA
The start codon has been indicated by bold capitals.

1/1 31/11
tca att ttg gta gct ctt aga aaa **ATG** gct ctg ctt att atc act gcc tta gtt ttt agt

61/21 91/31
agt tta tgg tca tct gtg tca cca gat gct caa ggg gat gca tta ttt gcg ttg agg agc

121/41 151/51
tcg tta cgt gca tct cct gaa cag ctt agt gat tgg aac cag aat caa gtc gat cct tgt

181/61 211/71
act tgg tct caa gtt att tgt gat gac aag aaa cat gtt act tct gta acc ttg tct tac

241/81 271/91
atg aac ttc tcc tcg gga aca ctg tct tca gga ata gga atc ttg aca act ctc aag act

301/101 331/111
ctt aca ttg aag gga aat gga ata atg ggt gga ata cca gaa tcc att gga aat ctg tct

361/121 391/131
agc ttg acc agc tta gat ttg gag gat aat cac tta act gat cgc att cca tcc act ctc

421/141 451/151
ggt aat ctc aag aat cta cag ttc ttt ttc aca gca aac aac ttg agc tgt ggt ggc act

481/161 511/171
ttc ccg caa cct tgt gta acc gag tcc agt cct tca ggt gat tca agc agt aga aaa act

541/181 571/191
gga atc atc gct gga gtt gtt agc gga ata gcg gtt att cta cta gga ttc ttc ttc ttt

601/201 631/211
ttc ttc tgc aag gat aaa cat aaa gga tat aaa cga gac gta ttt gtg gat gtt gca gga

661/221 691/231
acg aac ttt aaa aaa ggt ttg att tca ggt gaa gtg gac aga agg att gct ttt gga cag

721/241 751/251
ttg aga aga ttt gca tgg aga gag ctt cag ttg gct aca gat gag ttc agt gaa aag aat

781/261 811/271
gtt ctc gga caa gga ggc ttt ggg aaa gtt tac aaa gga ttg ctt tcg gat ggc acc aaa

841/281 871/291
gtc gct gta aaa aga ttg act gat ttt gaa cgt cca gga gga gat gaa gct ttc cag aga

901/301 931/311
gaa gtt gag atg ata agt gta gct gtt cat agg aat ctg ctt cgc ctt atc ggc ttt tgt

961/321 991/331
aca aca caa act gaa cga ctt ttg gtg tat cct ttc atg cag aat cta agt gtt gca tat

1021/341 1051/351
tgc tta aga gag att aaa ccc ggg gat cca gtt ctg gat tgg ttc agg agg aaa cag att

1081/361 1111/371
gcg tta ggt gca gca cga gga ctc gaa tat ctt cat gaa cat tgc aac ccg aag atc ata

1141/381 1171/391
cac aga gat gtg aaa gct gca aat gtg tta cta gat gaa gac ttt gaa gca gtg gtt ggt

1201/401 1231/411
gat ttt ggt tta gcc aag ttg gta gat gtt aga agg act aat gta acc act cag gtc cga

1261/421 1291/431
gga aca atg ggt cat att gca cca gaa tgt ata tcc aca ggg aaa tcg tca gag aaa acc

1321/441 1351/451
gat gtt ttc ggg tac gga att atg ctt ctg gag ctt gta act gga caa aga gca att gat

1381/461 1411/471

FIGURE 9a CONTD.

```
[illegible] tcg cgg tta gag gaa gaa gat gat gtc tta ttg cta gac cat gtg aag aaa ctg gaa

1441 481                                    1471/491
[illegible] aag aga tta gaa gac ata gta gat aag aag ctt gat gag gat tat ata aag gaa

1501 501                                    1531/511
gaa [illegible] gaa atg atg ata caa gta gct ctg cta tgc aca caa gca gca ccg gaa gaa cga

1561 521                                    1591/531
[illegible] gcg atg tcg gaa gta gta aga atg cta gaa gga gaa ggg ctt gca gag aga tgg gaa

1621 541                                    1651/551
[illegible] cag aat ctt gaa gtg acg aga caa gaa gag ttt cag agg ttg cag agg aga ttt

1681 561                                    1711/571
[illegible] ggt gaa gat tcc att aat aat caa gat gct att gaa tta tct ggt gga aga tag
```

Figure 9b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-2 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 3 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MALLIITALVFSSL
WSSVSPDAQG

DALFALRSSLR
ASPEQLSDWNQNQVD

PCTWSQVICDDKKHVTSV

TLSYMNFSSGTLSSGI
G ILTTLKTLTLKGNGIMGGI
PESIGNLSSLTSLDLEDNHLTDRI
PSTLGNLKNLQFFFTANNLSCGG

TFPQPCVTESSPSGDSSSRKTG

IIAGVVSGIAVIL
LGFFFFFFC

KDKHKGYKRDVFVDVAGTNFKKGLISGE
VDRRIAFGQLRRFAWRELQLAT

DEFSEKNVLGQGGFGKVYKGLLSD
GTKVAVKRLTDFERPGGDEAFQ
REVEMISVAVHRNLLRLIGFCT
TQTERLLVYPFMQNLSVAYCLR
EIKPGDPVLDWFRRKQIALGAA
RGLEYLHEHCNPKIIHRDVKAA
NVLLDEDFEAVVGDFGLAKLVD
VRRTNVTTQVRGTMGHIAPECI
STGKSSEKTDVFGYGIMLLELV
TGQRAIDFSRLEEEDDVLLLDH
VKKLEREKRLEDIVDKKLDEDY
IKEEVEMMIQVALLCTQAAPEE
RPAMSEVVRMLE

GEGLAERWEEWQNLEVTRQEEFQ

RLQRRFDWGEDSINNQDAIELSGGR

Figure 10a
Arabidopsis thaliana RKS3 cDNA
The start codon has been indicated by bold capitals.

1/1 31/11
aac ggt gaa agt ttc cat gat cct ctt cga gga ttc att caa aga aat tgc ttt aga tgg

61/21 91/31
aac aat cag aaa ttg atc tta caa tgt ttc **ATG** gcc tta gct ttt gtg gga atc act tcg

121/41 151/51
tca aca act caa cca gat atc gaa gga gga gct ctg ttg cag ctc aga gat tcg ctt aat

181/61 211/71
gat tcg agc aat cgt cta aaa tgg aca cgc gat ttt gtg agc cct tgc tat agt tgg tct

241/81 271/91
tat gtt acc tgc aga ggc cag agt gtt gtg gct cta aat ctt gcc tcg agt gga ttc aca

301/101 331/111
gga aca ctc tct cca gct att aca aaa ctg aag ttc ttg gtt acc tta gag tta cag aac

361/121 391/131
aat agt tta tct ggt gcc tta cca gat tct ctt ggg aac atg gtt aat cta cag act tta

421/141 451/151
aac cta tca gtg aat agt ttc agc gga tcg ata cca gcg agc tgg agt cag ctc tcg aat

481/161 511/171
cta aag cac ttg gat ctc tca tcc aat aat tta aca gga agc atc cca aca caa ttc ttc

541/181 571/191
tca atc cca aca ttc gat ttt tca gga act cag ctt ata tgc ggt aaa agt ttg aat cag

601/201 631/211
cct tgt tct tca agt tct cgt ctt cca gtc aca tcc tcc aag aaa aag ctg aga gac att

661/221 691/231
act ttg act gca agt tgt gtt gct tct ata atc tta ttc ctt gga gca atg gtt atg tat

721/241 751/251
cat cac cat cgc gtc cgc aga acc aaa tac gac atc ttt ttt gat gta gct ggg gaa gat

781/261 811/271
gac agg aag att tcc ttt gga caa cta aaa cga ttc tct tta cgt gaa atc cag ctc gca

841/281 871/291
aca gat agt ttc aac gag agc aat ttg ata gga caa gga gga ttt ggt aaa gta tac aga

901/301 931/311
ggt ttg ctt cca gac aaa aca aaa gtt gca gtg aaa cgc ctt gcg gat tac ttc agt cct

961/321 991/331
gga gga gaa gct gct ttc caa aga gag att cag ctc ata agc gtt gcg gtt cat aaa aat

1021/341 1051/351
ctc tta cgc ctt att ggc ttc tgc aca act tcc tct gag aga atc ctt gtt tat cca tac

1081/361 1111/371
atg gaa aat ctt agt gtt gca tat cga cta aga gat ttg aaa gcg gga gag gaa gga tta

1141/381 1171/391
gac tgg cca aca agg aag cgt gta gct ttt ggt tca gct cac ggt tta gag tat cta cac

1201/401 1231/411
gaa cat tgt aac ccg aag atc ata cac cgc gat ctc aag gct gca aac ata ctt tta gac

1261/421 1291/431
aac aat ttt gag cca gtt ctt gga gat ttc ggt tta gct aag ctt gtg gac aca tct ctg

1321/441 1351/451
act cat gtc aca act caa gtc cga ggc aca atg ggt cac att gcg cca gag tat ctc tgc

FIGURE 10a CONTD.

1381 461 1411/471
aaa gga aaa tca tct gaa aaa acc gat gtt ttt ggt tac ggt ata acg ctt ctt gag ctt

1441 481 1471/491
gtt act ggt cag cgc gca atc gat ttt tca cgc ttg gaa gaa gag gaa aat att ctc ttg

1501 501 1531/511
ctt gat cat ata aag aag ttg ctt aga gaa cag aga ctt aga gac att gtt gat agc aat

1561 521 1591/531
ttg act aca tat gac tcc aaa gaa gtt gaa aca atc gtt caa gtg gct ctt ctc tgc aca

1621 541 1651/551
caa ggc tca cca gaa gat aga cca gcg atg tct gaa gtg gtc aaa atg ctt caa ggg act

1681 561 1711/571
gtt ggt ttg gct gag aaa tgg act gaa tgg gaa caa ctt gaa gaa gtt agg aac aaa gaa

1741 581 1771/591
gta ttg ttg ctt ccg act tta ccg gct act tgg gat gaa gaa gaa acc acc gtt gat caa

1801 601
gaa tct atc cga tta tcg aca gca aga tga

Figure 10b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-3 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 3 leucine evenly residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MALAFVGITSSTTQPDIEG

GALLQLRDSLNDSSNRL
KWTRDFVS

PCYSWSYVTCRGQSVVAL

NLASSGFTGTLS
P AITKLKFLVTLELQNNSLSGAL
PDSLGNMVNLQTLNLSVNSFSGSI
PASWSQLSNLKHLDLSSNNLTGSI
PTQFFSIPTFEFSGTQLICGKS

LNQPCSSSRLPVTSSKKKLRD

ITLTASCVASIIL
FLGAMVMYHHH

RVRRTKYDIFFDVAGEDDR
KISFGQLKRFSLREIQLAT

DSFNESNLIGQGGFGKVYRGLLPD
KTKVAVKRLADYFSPGGEAAFQ
REIQLISVAVHKNLLRLIGFCT
TSSERILVYPYMENLSVAYRLR
DLKAGEEGLDWPTRKRVAFGSA
HGLEYLHEKCNPKIIHRDLKAA
NILLDNNFEPVLGDFGLAKLVD
TSLTHVTTQVRGTMGHIAPEYL
CTGKSSEKTDVFGYGITLLELV
TGQRAIDFSRLEEEENILLLD
HIKKLLREQRLRDIVDSNLTTY
DSKEVETIVQVALLCTQGSPED
RPAMSEVVKMLQ

GTGGLAEKWTEWEQLEEVRNKEALLL

PTLPATWDEEETTVDQESIRLSTAR

Figure 11a
Arabidopsis thaliana RKS4 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                                31/11
tct tcc ttc tcc ttc tgg taa tct aat cta aag ctt ttc ATG gtg gtg atg aag ata ttc

61/21                                              91/31
tct gtt ctg tta cta cta tgt ttc ttc gtt act tgt tct ctc tct tct gaa ccc aga aac

121/41                                             151/51
cct gaa gtc att aat ggt gac aaa ttc ttc atc ttt gtt ttg ttt ttt ccc aat tcc aga

181/61                                             211/71
gga gct cca agt cag tct ctt tca gga act tta tct ggg tct att gga aat ctc act aat

241/81                                             271/91
ctt cga caa gtg tca tta cag aac aat aac atc tcc ggt aaa atc cca ccg gag att tgt

301/101                                            331/111
tct ctt ccc aaa tta cag act ctg gat tta tcc aat aac cgg ttc tcc ggt gaa atc ccc

361/121                                            391/131
ggt tct gtt aac cag ctg agt aat ctc caa tat ctt gtt gct ggg aac cct ttg att tgt

421/141                                            451/151
aaa aac agc cta ccg gag att tgt tca gga tca atc agt gca agc cct ctt tct gtc tct

481/161                                            511/171
tta cgt tct tca tca gac aag caa gag gaa ggg tta ctt ggg ttg gga aat cta aga agc

541/181                                            571/191
ttc aca ttc agg gaa ctt cat gta gct acg gat ggt ttt agt tcc aag agt att ctt ggt

601/201                                            631/211
gct ggt ggg ttt ggt aat gtc tac aga gga aaa ttc ggg gat ggg aca gtg gtt gca gtg

661/221                                            691/231
aaa cga ttg aaa gat gtg aat gga acc tcc ggg aac tca cag ttt cgt act gag ctt gag

721/241                                            751/251
atg atc agc tta gct gtt cat agg aat ttg ctt cgg tta atc ggt tat tgt gcg agt tct

781/261                                            811/271
agc gaa aga ctt ctt gtt tac cct tac atg tcc aat ggc agc gtc gcc tct agg ctc aaa

841/281                                            871/291
gct aag cca gcg ttg gac tgg aac aca agg aag aag ata gcg att gga gct gca aga ggg

901/301                                            931/311
ttg ttt tat cta cac gag caa tgc gat ccc aag att att cac cga gat gtc aag gca gca

961/321                                            991/331
aac att ctc cta gat gag tat ttt gaa gca gtt gtt ggg gat ttt gga cta gca aag cta

1021/341                                           1051/351
ctc aac cac gag gat tca cat gtc aca acc gcg gtt aga gga act gtt ggt cac att gca

1081/361                                           1111/371
cct gag tat ctc tcc acc ggt cag tca tct gag aaa acc gat gtc ttt ggg ttc ggt ata

1141/381                                           1171/391
ctt ttg cta gag ctc atc aca gga atg aga gct ctc gag ttt ggc aag tct gtt agc cag

1201/401                                           1231/411
aaa gga gct atg cta gaa tgg gtg agg aag cta cac aag gaa atg aaa gta gag gag cta

1261/421                                           1291/431
gta gac cga gaa ctg ggg aca acc tac gat aga ata gaa gtt gga gag atg cta caa gtg

1321/441                                           1351/451
gca ctg ctc tgc act cag ttt ctt cca gct cac aga ccc aaa atg tct gaa gta gtt cag
```

FIGUUR 11a CONTD.

1381/461 1411/471
atg ctt gaa gga gat gga tta gct gag aga tgg gct gct tca cat gac cat tca cat ttc

1441/481 1471/491
tac cat gcc aac atg tct tac agg act att acc tct act gat ggc aac aac caa acc aaa

1501/501 1531/511
cat ctg ttt ggc tcc tca gga ttt gaa gat gaa gat gat aat caa gcg tta gat tca ttc

1561/521
gcc atg gaa cta tct ggt cca agg tag

Figure 11b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-4 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MVVMKLITMKIFSVLLLL
CFFVTCSLSSEPRNPEV

EALINIKNELHDP
HGVFKNWDEFSVD

PCSWTMISCSSDNLVIGL

GAPSQSLSGTLS
G SIGNLTNLRQVSLQNNNISGKI
PPEICSLPKLQTLDLSNNRFSGEI
PGSVNQLSNLQYLRLNNNSLSGPF
PASLSQIPHLSFLDLSYNNLRGPV
PKFPARTFNVAGNPLICKNS

LPEICSGSISASPL
SVSLRSSSGRRN

ILAVALGVSLGFAVSVIL
SLGFIWY

RKKQRRLTMLRISDKQEE
GLLGLGNLRSFTFRELHVAT

DGFSSKSILGAGGFGNVYRGKFGD
GTVVAVKRLKDVNGTSGNSQFR
TELEMISLAVHRNLLRLIGYCA
SSSERLLVYPYMSNGSVASRLK
AKPALDWNTRKKIAIGAA
RGLFYLHEQCDPKIIHRDVKAA
NILLDEYFEAVVGDFGLAKLLN
HEDSHVTTAVRGTVGHIAPEYL
STGQSSEKTDVFGFGILLLELI
TGMRALEFGKSVSQKGAMLEW
VRKLHKEMKVEELVDRELGTTY
DRIEVGEMLQVALLCTQFLPAH
RPKMSEVVQMLE

GDGLAERWAASHDHSHFYHANM
SYRTITSTDGNNQTKHLFG

SSGFEDEDDNQALDSFAMELSGPR

Figure 12a
Arabidopsis thaliana RKS5 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                     31/11
cta gag aat tct tat act ttt tct acg ATG gag att tct ttg atg aag ttt ctg ttt tta
 61/21                                   91/31
gga atc tgg gtt tat tat tac tct gtt ctt gac tct gtt tct gcc atg gat agt ctt tta

121/41                                  151/51
tct ccc aag ggt gtt aac tat gaa gtg gct gcg tta atg tca gtg aag aac aag atg aaa

181/61                                  211/71
gat gag aaa gag gtt ttg tct ggt tgg gat att aac tct gtt gat cct tgt act tgg aac

241/81                                  271/91
atg gtt ggt tgt tct tct gaa ggt ttt gtg gtt tct ctg tta ctt cag aat aat cag tta

301/101                                 331/111
act ggt ccg att cct tct gag tta ggc caa ctc tct gag ctt gaa acg ctt gat tta tcg

361/121                                 391/131
ggg aat cgg ttt agt ggt gaa atc cca gct tct tta ggg ttc tta act cac tta aac tac

421/141                                 451/151
ttg cgg ctt agc agg aat ctt tta tct ggg caa gtc cct cac ctc gtc gct ggc ctc tca

481/161                                 511/171
ggt ctt tct ttc ttg gat cta tct ttc aac aat cta agc gga cca act ccg aat ata tca

541/181                                 571/191
gca aaa gat tac agg att gta gga aat gca ttt ctt tgt ggt cca gct tcc caa gag ctt

601/201                                 631/211
tgc tca gat gct aca cct gtg aga aat gtg cag caa gac tac gaa ttt gaa atc ggc cat

661/221                                 691/231
ctg aaa agg ttc agt ttt cgc gaa ata caa acc gca aca agc aat ttt agt cca aag aac

721/241                                 751/251
att ttg gga caa gga ggg ttt ggg atg gtt tat aaa ggg tat ctc cca aat gga act gtg

781/261                                 811/271
gtg gca gtt aaa aga ttg aaa gat ccg att tat aca gga gaa gtt cag ttt caa acc gaa

841/281                                 871/291
gta gag atg att ggc tta gct gtt cac cgt aac ctt tta cgc ctc ttt gga ttc tgt atg

901/301                                 931/311
acc ccg gaa gag aga atg ctt gtg tat ccg tac atg cca aat gga agc gta gct gat cgt

961/321                                 991/331
ctg aga gat tgg aat cgg agg ata agc att gca ctc ggc gca gct cga gga ctt gtt tac

1021/341                                1051/351
ttg cac gag caa tgc aat cca aag att att cac aga gac gtc aaa gct gca aat att cta

1081/361                                1111/371
ctt gat gag agc ttt gaa gca ata gtt ggc gat ttt ggt cta gca aag ctt tta gac cag

1141/381                                1171/391
aga gat tca cat gtc act acc gca gtc cga gga acc att gga cac atc gct ccc gag tac

1201/401                                1231/411
ctt tcc act gga cag tcc tca gag aaa acc gat gtt ttc gga ttc gga gta cta atc ctt

1261/421                                1291/431
gaa ctc ata aca ggt cat aag atg att gat caa ggc aat ggt caa gtt cga aaa gga atg

1321/441                                1351/451
ata ttg agc tgg gta agg aca ttg aaa gca gag aag aga ttt gca gag atg gtg gac aga

1381/461                                1411/471
gat ttg aag gga gag ttt gat gat ttg gtg ttg gag gaa gta gtg gaa ttg gct ttg ctt
```

FIGUUR 12a CONTD.

1441/481 1471/491
tgt aca cag cca cat ccg aat cta aga ccg agg atg tct caa gtg ttg aag gta cta gaa

1501/501 1531/511
ggt tta gtg gaa cag tgt gaa gga ggg tat gaa gct aga gct cca agt gtc tct agg aac

1561/521 1591/531
tac agt aat ggt cat gaa gag cag tcc ttt att att gaa gcc att gag ctc tct gga cca

1621/541
cga tga tag

Figure 12b

Predicted amino acid sequence of the Arabidopsis thaliana RKS-5 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues.
The fifth domain has no clear function.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MEISLMKFLFLGIWVYYY
SVLDSVSAM

DSLLSPKWAALMSVKNKMKDE
KEVLSGWDINSVD

PCTWNMVGCSSEGFVVS

LLLQNNQLTGPI
PSELGQLSELETLDLSGNRFSGEI
PASLGFLTHLNYLRLSRNLLSGQV
PHLVAGLSGLSFLDLSFNNLSGPT
P NISAKDYRIVGNAFLCGPA

SQELCSDATPVRNGMLLRKFFAKLYL
KHGFVYLTSCNRSAATGLSEKDNSK

HHSLVLSFAFGIVVA
FIISLMFLFFWVLWH

RSRLSRSHGTYLIVSLCLSYTIYVKTLLKSA
LLFMDFLVQQDYEFEIGHLKRFSFREIQTAT

SNFSPKNILGQGGFGMVYKGYLPN
GTVVAVKRLKDPIYTGEVQFQ
TEVEMIGLAVHRNLLRLFGFCM
TPEERMLVYPYMPNGSVADRLR
DWNRRISIALGAA
RGLVYLHEQCNPKIIHRDVKAA
NILLDESFEAIVGDFGLAKLLD
QRDSHVTTAVRGTIGHIAPEYL
STGQSSEKTDVFGFGVLILELI
TGHKMIDQGNGQVRKGMILSW
VRTLKAEKRFAEMVDRDLKGEF
DDLVLEEVVELALLCTQPHPNL
RPRMSQVLKV

LEGLVEQCEGGYEARA

PASVSRNYSNGHEEQSFIIEAIELSGPR

Figure 13a
Arabidopsis thaliana RKS6 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                                 31/11
ATT GTT TCC TTC TTT TGG GAT TTT CTC CTT GGA TGG AAC CAG CTC AAT TAA TGA GAT GAG

61/21                                               91/31
ATG AGA ATG TTC AGC TTG CAG AAG ATG GCT ATG GCT TTT ACT CTC TTG TTT TTT GCC TGT
121/41
                                        151/51
TTA TGC TCA TTT GTG TCT CCA GAT GCT CAA GGG GAT GCA CTG TTT GCG TTG AGG ATC TCC

181/61                                              211/71
TTA CGT GCA TTA CCG AAT CAG CTA AGT GAC TGG AAT CAG AAC CAA GTT AAT CCT TGC ACT

241/81                                              271/91
TGG TCC CAA GTT ATT TGT GAT GAC AAA AAC TTT GTC ACT TCT CTT ACA TTG TCA GAT ATG

301/101                                             331/111
AAC TTC TCG GGA ACC TTG TCT TCA AGA GTA GGA ATC CTA GAA AAT CTC AAG ACT CTT ACT

361/121                                             391/131
TTA AAG GGA AAT GGA ATT ACG GGT GAA ATA CCA GAA GAC TTT GGA AAT CTG ACT AGC TTG

421/141                                             451/151
ACT AGT TTG GAT TTG GAG GAC AAT CAG CTA ACT GGT CGT ATA CCA TCC ACT ATC GGT AAT

481/161                                             511/171
CTC AAG AAA CTT CAG TTC TTG ACC TTG AGT AGG AAC AAA CTT AAT GGG ACT ATT CCG GAG

541/181                                             571/191
TCA CTC ACT GGT CTT CCA AAC CTG TTA AAC CTG CTG CTT GAT TCC AAT AGT CTC AGT GGT

601/201                                             631/211
CAG ATT CCT CAA AGT CTG TTT GAG ATC CCA AAA TAT AAT TTC ACG TCA AAC AAC TTG AAT

661/221                                             691/231
TGT GGC GGT CGT CAA CCT CAC CCT TGT GTA TCC GCG GTT GCC CAT TCA GGT GAT TCA AGC

721/241                                             751/251
AAG CCT AAA ACT GGC ATT ATT GCT GGA GTT GTT GCT GGA GTT ACA GTT GTT CTC TTT GGA

781/261                                             811/271
ATC TTG TTG TTT CTG TTC TGC AAG GAT AGG CAT AAA GGA TAT AGA CGT GAT GTG TTT GTG

841/281                                             871/291
GAT GTT GCA GGT GAA GTG GAC AGG AGA ATT GCA TTT GGA CAG TTG AAA AGG TTT GCA TGG

901/301                                             931/311
AGA GAG CTC CAG TTA GCG ACA GAT AAC TTC AGC GAA AAG AAT GTA CTT GGT CAA GGA GGC

961/321                                             991/331
TTT GGG AAA GTT TAC AAA GGA GTG CTT CCG GAT ACA CCC AAA GTT GCT GTG AAG AGA TTG

1021/341                                            1051/351
ACG GAT TTC GAA AGT CCT GGT GGA GAT GCT GCT TTC CAA AGG GAA GTA GAG ATG ATA AGT

1081/361                                            1111/371
GTA GCT GTT CAT AGG AAT CTA CTC CGT CTT ATC GGG TTC TGC ACC ACA CAA ACA GAA CGC

1141/381                                            1171/391
CTT TTG GTT TAT CCC TTC ATG CAG AAT CTA AGT CTT GCA CAT CGT CTG AGA GAG ATC AAA

1201/401                                            1231/411
GCA GGC GAC CCG GTT CTA GAT TGG GAG ACG AGG AAA CGG ATT GCC TTA GGA GCA GCG CGT

1261/421                                            1291/431
GGT TTT GAG TAT CTT CAT GAA CAT TGC AAT CCG AAG ATC ATA CAT CGT GAT GTG AAA GCA

1321/441                                            1351/451
GCT AAT GTG TTA CTA GAT GAA GAT TTT GAA GCA GTG GTT GGT GAT TTT GGT TTA GCC AAG

1381/461                                            1411/471
```

FIGURE 13a CONTD.

CTA GTA GAT GTT AGA AGG ACT AAT GTG ACT ACT CAA GTT CGA GGA ACA ATG GGT CAC ATT

1441 481 1471/491
GCA CCA GAA TAT TTA TCA ACA GGG AAA TCA TCA GAG AGA ACC GAT GTT TTC GGG TAT GGA

1501 501 1531/511
ATT ATG CTT CTT GAG CTT GTT ACA GGA CAA CGC GCA ATA GAC TTT TCA CGT TTG GAG GAA

1561 521 1591/531
GAA GAT GAT GTC TTG TTA CTT GAC CAC GTG AAG AAA CTG GAA AGA GAG AAG AGA TTA GGA

1621 541 1651/551
GTA ATC GTA GAT AAG AAT TTG GAT GGA GAG TAT ATA AAA GAA GAA GTA GAG ATG ATG ATA
A

1681 561 1711/571
AAA GTG GCT TTG CTT TGT ACA CAA GGT TCA CCA GAA GAC CGA CCA GTG ATG TCT GAA GTT

1741 581 1771/591
GTC AAG ATG TTA GAA GGA GAA GGG CTT GCG GAG AGA TGG GAA GAG TGG CAA AAC GTG GAA

1801 601 1831/611
GTT ACG AGA CGT CAT GAG TTT GAA CGG TTG CAG AGG AGA TTT GAT TGG GGT GAA GAT TCT

1861 621 1891/631
ATC CAT AAC CAA GAT GCC ATT GAA TTA TCT GGT GGA AGA TGA CCA AAA ACA TCA AAC CTT

Figure 13b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-6 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MRMFSL
QKMAMAFTLLFFACLCSFVSPDAQG

DALFALRISLRALP
NQLSDWNQNQVN

PCTWSQVICDDKNFVTSL

TLSDMNFSGTLSSRV
GILENLKTLTLKGNGITGEI
PEDFGNLTSLTSLDLEDNQLTGRI
PSTIGNLKKLQFLTLSRNKLNGTI
PESLTGLPNLLNLLLDSNSLSGQI
PQSLFEIPKYNFTSNNLNCGG

RQPHPCVSAVAHSGDSSKPKTG

IIAGVVAGVTVVL
FGILLFLFC

KDRHKGYRRDVFVDVAGE
VDRRIAFGQLKRFAWRELQLAT

DNFSEKNVLGQGGFGKVYKGVLPD
TPKVAVKRLTDFESPGGDAAFQ
REVEMISVAVHRNLLRLIGFCT
TQTERLLVYPFMQNLSLAHRLR
EIKAGDPVLDWETRKRIALGAA
RGFEYLHEHCNPKIIHRDVKAA
NVLLDEDFEAVVGDFGLAKLVD
VRRTNVTTQVRGTMGHIAPEYL
STGKSSERTDVFGYGIMLLELV
TGQRAIDFSRLEEEDDVLLLDH
VKKLEREKRLGAIVDKNLDGEY
IKEEVEMMIQVALLCTQGSPED
RPVMSEVVRMLE

GEGLAERWEEWQNVEVTRRHEFE

RLQRRFDWGEDSMHNQDAIELSGGR

FIGURE 14a

Arabidopsis thaliana RKS8 cDNA
The start codon has been indicated by bold capitals.

1/1 31/11
GTT TTT TTT TTT TTA CCC TCT TGG AGG ATC TGG GAG GAG AAA TTT GCT TTT TTT TGG TAA

61/21 91/31
ATG GGG AGA AAA AAG TTT GAA GCT TTT GGT TTT GTC TGC TTA ATC TCA CTG CTT CTT CTG

121/41 151/51
TTT AAT TCG TTA TGG CTT GCC TCT TCT AAC ATG GAA GGT GAT GCA CTG CAC AGT TTG AGA

181/61 211/71
GCT AAT CTA GTT GAT CCA AAT AAT GTC TTG CAA AGC TGG GAT CCT ACG CTT GTT AAT CCG

241/81 271/91
TGT ACT TGG TTT CAC GTA ACG TGT AAC AAC GAG AAC AGT GTT ATA AGA GTC GAT CTT GGG

301/101 331/111
AAT GCA GAC TTG TCT GGT CAG TTG GTT CCT CAG CTA GGT CAG CTC AAG AAC TTG CAG TAC

361/121 391/131
TTG GAG CTT TAT AGT AAT AAC ATA ACC GGG CCG GTT CCA AGC GAT CTT GGG AAT CTG ACA

421/141 451/151
AAC TTA GTG AGC TTG GAT CTT TAC TTG AAC AGC TTC ACT GGT CCA ATT CCA GAT TCT CTA

481/161 511/171
GGA AAG CTA TTC AAG CTT CGC TTT CTT CGG CTC AAC AAT AAC AGT CTC ACC GGA CCA ATT

541/181 571/191
CCC ATG TCA TTG ACT AAT ATC ATG ACC CTT CAA GTT TTG GAT CTG TCG AAC AAC CGA TTA

601/201 631/211
TCC GGA TCT GTT CCT GAT AAT GGT TCC TTC TCG CTC TTC ACT CCC ATC AGT TTT GCT AAC

661/221 691/231
AAC TTG GAT CTA TGC GGC CCA GTT ACT AGC CGT CCT TGT CCT GGA TCT CCC CCG TTT TCT

721/241 751/251
CCT CCA CCA CCT TTT ATA CCA CCT CCC ATA GTT CCT ACA CCA GGT GGG TAT AGT GCT ACT

781/261 811/271
GGA GCC ATT GCG GGA GGA GTT GCT GCT GGT GCT GCT TTA CTA TTT GCT GCC CCT GCT TTA

841/281 871/291
GCT TTT GCT TGG TGG CGT AGA AGA AAA CCT CAA GAA TTC TTC TTT GAT GTT CCT GCC GAA

901/301 931/311
GAG GAC CCT GAG GTT CAC TTG GGG CAG CTT AAG CGG TTC TCT CTA CGG GAA CTT CAA GTA

961/321 991/331
GCA ACT GAT AGC TTC AGC AAC AAG AAC ATT TTG GGC CGA GGT GGG TTC GGA AAA GTC TAC

1021/341 1051/351
AAA GGC CGT CTT GCT GAT GGA ACA CTT GTT GCA GTC AAA CGG CTT AAA GAA GAG CGA ACC

1081/361 1111/371
CCA GGT GGC GAG CTC CAG TTT CAG ACA GAA GTG GAG ATG ATA AGC ATG GCC GTT CAC AGA

1141/381 1171/391
AAT CTC CTC AGG CTA CGC GGT TTC TGT ATG ACC CCT ACC GAG AGA TTG CTT GTT TAT CCT

1201/401 1231/411
TAC ATG GCT AAT GGA AGT GTC GCT TCC TGT TTG AGA GAA CGT CCA CCA TCA CAG TTG CCT

1261/421 1291/431
CTA GCC TGG TCA ATA AGA CAG CAA ATC GCG CTA GGA TCA GCG AGG GGT TTG TCT TAT CTT

1321/441 1351/451
CAT GAT CAT TGC GAC CCC AAA ATT ATT CAC CGT GAT GTG AAA GCT GCT AAT ATT CTG TTG

FIGURE 14a, CONTD.

1381/461 1411/471
GAC GAG GAA TTT GAG GCG GTG GTA GGT GAT TTC GGG TTA GCT AGA CTT ATG GAC TAT AAA

1441/481 1471/491
GAT ACT CAT GTC ACA ACG GCT GTG CGT GGG ACT ATT GGA CAC ATT GCT CCT GAG TAT CTC

1501/501 1531/511
TCA ACT GGA AAA TCT TCA GAG AAA ACT GAT GTT TTT GGC TAC GGG ATC ATG CTT TTG GAA

1561/521 1591/531
CTG ATT ACA GGT CAG AGA GCT TTT GAT CTT GCA AGA CTG GCG AAT GAC GAT GAC GTT ATG

1621/541 1651/551
CTC CTA GAT TGG GTG AAA GGG CTT TTG AAG GAG AAG AAG CTG GAG ATG CTT GTG GAT CCT

1681 561 1711/571
GAC CTG CAA AGC AAT TAC ACA GAA GCA GAA GTA GAA CAG CTC ATA CAA GTG GCT CTT CTC

1741 581 1771/591
TGC ACA CAG AGC TCA CCT ATG GAA CGA CCT AAG ATG TCT GAG GTT GTT CGA ATG CTT GAA

1801/601 1831/611
GTT GAC GGT TTA GCG GAG AAA TGG GAC GAG TGG CAG AAA GTG GAA GTT CTC AGG CAA GAA

1861/621 1891/631
CTG GAG CTC TCT TCT CAC CCC ACC TCT GAC TGG ATC CTT GAT TCG ACT GAT AAT CTT CAT

1921/641
GCT ATG GAG TTG TCT GGT CCA AGA TAA AC

Figure 14b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-8 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 4 leucine evenly spaced residues, each seperated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 5 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MGRKKFEAFGFVCLISLLLLFNSL
WLASSNMEG

DALHSLRANLVDP
NNVLQSWDPTLVN

PCTWFHVTCNNENSVIRV

DLGNADLSGQLV
P QLGQLKNLQYLELYSNNITGPV
PSDLGNLTNLVSLDLYLNSFTGPI
PDSLGKLFKLRFLRLNNNSLTGPI
PMSLTNIMTLQVLDLSNNRLSGSV
PDNGSFSLFTPISFANNLDLCGPV

TLRPCPGSPPFSPPPP
FIPPPIVPTPGGYSATG

AIAGGVAAGAAL
LFAAPALAFAWW

RRRKPQEFFFDVPAEEDPE
VHLGQLKRFSLRELQVAT

DSFSNKNILGRGGFGKVYKGRLAD
GTLVAVKRLKEERTPGGELQFQ
TEVEMISMAVHRNLLRLRGFCM
TPTERLLVYPYMANGSVASCLR
ERPPSQLPLAWSIRQQIALGSA
RGLSYLHDHCDPKIIHRDVKAA
NILLDEEFEAVVGDFGLARLMD
YKDTHVTTAVRGTIGHIAPEYL
STGKSSEKTDVFGYGIMLLELI
TGQRAFDLARLANDDDVMLLDW
VKGLLKEKKLEMLVDPDLQSNY
TEAEVEQLIQVALLCTQSSPME
RPKMSEVVRMLE

GDGLAEKWDEWQKVEVLRQEVELS

SHPTSDWILDSTDNLHAMELSGPR

Figure 15a
Arabidopsis thaliana RKS10 cDNA
The start codon has been indicated by bold capitals.

1/1 31/11
atc agg ggt ttt aac aat gat gga ttt tct ctg atg agg gat agt tct agg gtt tgt ttt

61/21 91/31
taa tct ctt gag gat aaa **ATG** gaa cga aga tta atg atc cct tgc ttc ttt tgg ttg att

121/41 151/51
ctc gtt ttg gat ttg gtt ctc aga gtc tcg ggc aac gcc gaa ggt gat gct cta agt gca

181/61 211/71
ctg aaa aac agt tta gcc gac cct aat aag gtg ctt caa agt tgg gat gct act ctt gtt

241/81 271/91
act cca tgt aca tgg ttt cat gtt act tgc aat agc gac aat agt gtt aca cgt gtt gac

301/101 331/111
ctt ggg aat gca aat cta tct gga cag ctc gta atg caa ctt ggt cag ctt cca aac ttg

361/121 391/131
cag tac ttg gag ctt tat agc aat aac att act ggg aca atc cca gaa cag ctt gga aat

421/141 451/151
ctg acg gaa ttg gtg agc ttg gat ctt tac ttg aac aat tta agc ggg cct att cca tca

481/161 511/171
act ctc ggc cga ctt aag aaa ctc cgt ttc ttg cgt ctt aat aac aat agc tta tct gga

541/181 571/191
gaa att cca agg tct ttg act gct gtc ctg acg cta caa gtt ctt ttt gcc aac acc aag

601/201 631/211
ttg act ccc ctt cct gca tct cca ccg cct cct atc tct cct aca ccg cca tca cct gca

661/221 691/231
ggg agt aat aga att act gga gcg att gcg gga gga gtt gct gca ggt gct gca ctt cta

721/241 751/251
ttt gct gtt ccg gcc att gca cta gct tgg tgg cga agg aaa aag ccg cag gac cac ttc

781/261 811/271
ttt gat gta cca gct gaa gag gac cca gaa gtt cat tta gga caa ctg aag agg ttt tca

841/281 871/291
ttg cgt gaa cta caa gtt gct tcg gat aat ttt agc aac aag aac ata ttg ggt aga ggt

901/301 931/311
ggt ttt ggt aaa gtt tat aaa gga cgg tta gct gat ggt act tta gtg gcc gtt aaa agg

961/321 991/331
cta aaa gag gag cgc acc caa ggt ggc gaa ctg cag ttc cag aca gag gtt gag atg att

1021/341 1051/351
agt atg gcg gtt cac aga aac ttg ctt cgg ctt cgt gga ttt tgc atg act cca acc gaa

1081/361 1111/371
aga ttg ctt gtt tat ccc tac atg gct aat gga agt gtt gcc tcc tgt tta aga gaa cgt

1141/381 1171/391
ccc gag tcc cag cca cca ctt gat tgg cca aag aga cag cgt att gcg ttg gga tct gca

1201/401 1231/411
aga ggg ctt gcg tat tta cat gat cat tgc gac cca aag att att cat cga gat gtg aaa

1261/421 1291/431
gct gca aat att ttg ttg gat gaa gag ttt gaa gcc gtg gtt ggg gat ttt gga ctt gca

1321/441 1351/451
aaa ctc atg gac tac aaa gac aca cat gtg aca acc gca gtg cgt ggg aca att ggt cat

1381/461 1411/471

FIGUUR 15a CONTD.

ata gcc cct gag tac ctt tcc act gga aaa tca tca gag aaa acc gat gtc ttt ggg tat

1441/481 1471/491
gga gtc atg ctt ctt gag ctt atc act gga caa agg gct ttt gat ctt gct cgc ctc gcg

1501/501 1531/511
aat gat gat gat gtc atg tta cta gac tgg gtg aaa ggg ttg tta aaa gag aag aaa ttg

1561/521 1591/531
gaa gca cta gta gat gtt gat ctt cag ggt aat tac aaa gac gaa gaa gtg gag cag cta

1621/541 1651/551
atc caa gtg gct tta ctc tgc act cag agt tca cca atg gaa aga ccc aaa atg tct gaa

1681/561 1711/571
gtt gta aga atg ctt gaa gga gat ggt tta gct gag aga tgg gaa gag tgg caa aag gag

1741/581 1771/591
gaa atg ttc aga caa gat ttc aac tac cca acc cac cat cca gcc gtg tct ggc tgg atc

1801/601 1831/611
att ggc gat tcc act tcc cag atc gaa aac gaa tac ccc tcg ggt cca aga taa gat tcg

1861/621 1891/631
aaa cac gaa tgt ttt ttc tgt att ttg ttt ttc tct gta ttt att gag ggt ttt agc ttc

Figure 15b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-10 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 4 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MERRLMIPCFFWLILVL
DLVLRVSGNAEG

DALSALKNSLADP
NKVLQSWDATLVT

PCTWFHVTCNSDNSVTRV

DLGNANLSGQLV
M QLGQLPNLQYLELYSNNITGTI
PEQLGNLTELVSLDLYLNNLSGPI
PSTLGRLKKLRFLRLNNNSLSGEI
PRSLTAVLTLQVLFANTK LTPL

PASPPPPISPTPPSPAGSNRITG

AIAGGVAAGAAL
LFAVPAIALAWW

RRKKPQDHFFDVPAEEDPE
VHLGQLKRFSLRELQVAS

DNFSNKNILGRGGFGKVYKGRLAD
GTLVAVKRLKEERTQGGELQFQ
TEVEMISMAVHRNLLRLRGFCM
TPTERLLVYPYMANGSVASCLR
ERPESQPPLDWPKRQRIALGSA
RGLAYLHDHCDPKIIHRDVKAA
NILLDEEFEAVVGDFGLAKLMD
YKDTHVTTAVRGTIGHIAPEYL
STGKSSEKTDVFGYGVMLLELI
TGQRAFDLARLANDDDVMLLDW
VKGLLKEKKLEALVDVDLQGNY
KDEEVEQLIQVALLCTQSSPME
RPKMSEVVRMLE

GDGLAERWEEWQKEEMFRQDFNYPTHH

PAVSGWIIGDSTSQIENEYPSGPR

Figure 16a
Arabidopsis thaliana RKS11 cDNA
The start codon has been indicated by bold capitals.

tgttaacctctcgtaactaaaatcttcc
ATGaagattcaaattcatctcctttactcgttcttgttcctctgtttctctactctcact
ctatctttctgagcccagaaaccctgaagttgaggcgttgataagtataaggaacaatttg
catgatcctcatggagctttgaacaattgggacgagttttcagttgatccttgtagctgg
gctatgatcacttgctctcccgacaacctcgtcattggactgtcattgcaaaataacaac
atctccggcaaaattccaccggagctcggttttctacccaaattacaaaccttggatctt
tccaacaaccgattctccggtgacatccctgtttccatcgaccagctaagcagccttcaa
tatctgagactcaacaacaactctttgtctgggcccttccctgcttctttgtcccaaatt
Cctcacctctccttcttggacttgtcttacaacaatctcagtggccctgttcctaaattc
ccagcaaggactttcaa
cgttgctggtaatcctttgatttgtagaagcaacccacctgagatttgttctgga
tcaatcaatgcaagtccactttctgtttctttgagctcttcatcagcagataaacaagag
gaagggcttcaaggacttgggaatctaagaagcttcacattcagagaactccatgtttat
acagatggtttcagttccaagaacattctcggcgctggtggattcggtaatgtgtacaga
ggcaagcttggagatgggacaatggtggcagtgaaacggttgaaggatattaatggaacc
tcaggggattcacagtttcgtatggagctagagatgattagcttagctgttcataagaat
ctgcttcggttaattggttattgcgcaacttctggtgaaaggcttcttgtttacccttac
atgcctaatggaagcgtcgcctctaagcttaaatctaaaccggcattggactggaacatg
aggaagaggatagcaattggtgcagcgagaggtttgttgtatctacatgagcaatgtgat
cccaagatcattcatagagatgtaaaggcagctaatattctcttagacgagtgctttgaa
gctgttgttggtgactttggactcgcaaagctccttaaccatgcggattctcatgtcaca
actgcggtccgtggtacggttggccacattgcacctgaatatctctccactggtcagtct
tctgagaaaaccgatgtgtttgggttcggtatactattgctcgagctcataaccggactg
agagctcttgagtttggtaaaaccgttagccagaaaggagctatgcttgaatgggtgagg
aaattacatgaagagatgaaagtagaggaactattggatcgagaactcggaactaactac
gataagattgaagttggagagatgttgcaagtggctttgctatgcacacaatatctgcca
gctcatcgtcctaaaatgtctgaagttgttttgatgcttgaaggcgatggattagccgag
agatgggctgcttcgcataaccattcacatttctaccatgccaatatctctttcaagaca
atctcttctctgtctactacttctgtctcaaggcttgacgcacattgcaatgatccaact
tatcaaatgtttggatcttcggctttcgatgatgacgatgatcatcagcctttagattcc
tttgccatggaactatccggtccaagataacacaatgaaaaaaaaaaaaaaaaaaaaaa

Figure 16b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-11 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 3 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 3 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MKIQIKLLYSFLFLCFSTL
TLSSEPRNPEV

EALISIRNNLHDP
HGALNNWDEFSVD

PCSWAMITCSPDNLVIGL

SLQNNNISGKI
PPELGFLPKLQTL DLSNNRFSGDI
PVSIDQLSSLQYLDLSYNNLSGPV
PKFPARTFNVAGNPLICRSN

PPEICSGSINASPL
SVSLSSSSGTRSNR

LAIALSVSLGSVVILVLALGSFCWY

RKKQRRLLILNLNADKQEE
GLQGLGNLRSFTFRELHVYT

DGFSSKNILGAGGFGNVYRGKLGD
GTMVAVKRLKDINGTSGDSQFR
MELEMISLAVHKNLLRLIGYCA
TSGERLLVYPYMPNGSVASKLK
SKPALDWNMRKRIAIGAA
RGLLYLHEQCDPKIIHRDVKAA
NILLDECFEAVVGDFGLAKLLN
HADSHVTTAVRGTVGHIAPEYL
STGQSSEKTDVFGFGILLLELI
TGLRALEFGKTVSQKGAMLEW
VRKLHEEMKVEELLDRELGTNY
DKIEVGEMLQVALLCTQYLPAH
RPKMSEVVLMLE

GDGLAERWAASHNHSHFYHANISFKT
ISSLSTTSVSRLDAHCND

PTYQMFGSSAFDDDDDHQPLDSFAMELSGPR

Figure 17a
Arabidopsis thaliana RKS12 cDNA
The start codon has been indicated by bold capitals.

1/1 31/11
ttt aaa aac ctt gct agt tct caa ttc tca tga ctt tgc ttt tag tct tag aag tgg aaa

61/21 91/31
ATG gaa cat gga tca tcc cgt ggc ttt att tgg ctg att cta ttt ctc gat ttt gtt tcc

121/41 151/51
aga gtc acc gga aaa aca caa gtt gat gct ctc att gct cta aga agc agt tta tca tca

181/61 211/71
ggt gac cat aca aac aat ata ctc caa agc tgg aat gcc act cac gtt act cca tgt tca

241/81 271/91
tgg ttt cat gtt act tgc aat act gaa aac agt gtt act cgt ctg gaa ctt ttt aac aat

301/101 331/111
aat att act ggg gag ata cct gag gag ctt ggc gac ttg atg gaa cta gta agc ttg gac

361/121 391/131
ctt ttt gca aac aac ata agc ggt ccc atc cct tcc tct ctt ggc aaa cta gga aaa ctc

421/141 451/151
cgc ttc ttg cgt ctt tat aac aac agc tta tct gga gaa att cca agg tct ttg act gct

481/161 511/171
ctg ccg ctg gat gtt ctt gat atc tca aac aat cgg ctc agt gga gat att cct gtt aat

541/181 571/191
ggt tcc ttt tcg cag ttc act tct atg agt ttt gcc aat aat aaa tta agg ccg cga cct

601/201 631/211
gca tct cct tca cca tca cct tca gga acg tct gca gca ata gta gtg gga gtt gct gcg

661/221 691/231
ggt gca gca ctt cta ttt gcg ctt gct tgg tgg ctg aga aga aaa ctg cag ggt cac ttt

721/241 751/251
ctt gat gta cct gct gaa gaa gac cca gag gtt tat tta gga caa ttt aaa agg ttc tcc

781/261 811/271
ttg cgt gaa ctg cta gtt gct aca gag aaa ttt agc aaa aga aat gta ttg ggc aaa gga

841/281 871/291
cgt ttt ggt ata ttg tat aaa gga cgt tta gct gat gac act cta gtg gct gtg aaa cgg

901/301 931/311
cta aat gaa gaa cgt acc aag ggt ggg gaa ctg cag ttt caa acc gaa gtt gag atg atc

961/321 991/331
agt atg gcc gtt cat agg aac ttg ctt cgg ctt cgt ggc ttt tgc atg act cca act gaa

1021/341 1051/351
aga tta ctt gtt tat ccc tac atg gct aat gga agt gtt gct tct tgt tta aga gag cgt

1081/361 1111/371
cct gaa ggc aat cca gcc ctt gac tgg cca aaa aga aag cat att gct ctg gga tca gca

1141/381 1171/391
agg ggg ctc gca tat tta cac gat cat tgc gac caa aag atc att cac ctg gat gtg aaa

1201/401 1231/411
gct gca aat ata ctg tta gat gaa gag ttt gaa gct gtt gtt gga gat ttt ggg cta gca

1261/421 1291/431
aaa tta atg aat tat aac gac tcc cat gtg aca act gct gta cgg ggt acg att ggc cat

1321/441 1351/451
ata gcg ccc gag tac ctc tcg aca gga aaa tct tct gag aag act gat gtt ttt ggg tac

FIGUUR 17a CONTD.

1381/461 1411/471
ggg gtc atg ctt ctc gag ctc atc act gga caa aag gct ttc gat ctt gct cgg ctt gca

1441/481 1471/491
aat gat gat gat atc atg tta ctc gac tgg gtg aaa gag gtt ttg aaa gag aag aag ttg

1501/501 1531/511
gaa agc ctt gtg gat gca gaa ctc gaa gga aag tac gtg gaa aca gaa gtg gag cag ctg

1561/521 1591/531
ata caa atg gct ctg ctc tgc act caa agt tct gca atg gaa cgt cca aag atg tca gaa

1621/541 1651/551
gta gtg aga atg ctg gaa gga gat ggt tta gct gag aga tgg gaa gaa tgg caa aag gag

1681/561 1711/571
gag atg cca ata cat gat ttt aac tat caa gcc tat cct cat gct ggc act gac tgg ctc

1741/581 1771/591
atc ccc tat tcc aat tcc ctt atc gaa aac gat tac ccc tcg ggg cca aga taa cct ttt

1801/601 1831/611
aga aag ggt cat ttc ttg tgg gtt ctt caa caa gta tat ata tag gta gtg aag ttg taa

1861/621 1891/631
gaa gca aaa ccc cac att cac ctt tga ata tca cta ctc tat aaaaaaaaaaaaaaaaaaaaaa

Figure 17b

Predicted amino acid sequence of the Arabidopsis thaliana RKS-12 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).

At the predicted extracellular domain the first domain represents a signal sequence.

The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.

The third domain contains conserved cysteine residues, involved in disulphate bridge formation.

The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues.

The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.

The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.

The seventh domain has an unknown function.

The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.

The ninth domain has an unknown function.

The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

[illegible]
[illegible]TGKTQV

[illegible]SGDHTNNILQ
[illegible]

[illegible]NSVTRL

ELFNNNITGEI
[illegible]LVSLDLFANNISGPI
[illegible]LRLYNNSLSGEI
[illegible]LDISNNRLSGDI
[illegible]RFA NNKLRPR

[illegible]

[illegible]AALLFALAWWL

[illegible]PAAEEDPE
[illegible]LRELLVAT

[illegible]GRFGILYKGRLAD
[illegible]ERTKGGELQFQ
[illegible]RNLLRLRGFCM
[illegible]ANGSVASCLR
[illegible]PKRKHIALGSA
[illegible]IIHLDVKAA
[illegible]VGDFGLAKLMN
[illegible]TIGHIAPEYL
[illegible]TGYGVMLLELI
[illegible]ANDDDIMLLDW
[illegible]LVDAELEGKY
[illegible]ALLCTQSSAME
[illegible]

[illegible]KEEMPIHDFNYQAY

[illegible]SLIENDYPSGPR

Figure 18a
Arabidopsis thaliana RKS13 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                                 31/11
taa taa acc tct aat aat aat ggc ttt gct ttt act ctg ATG aca agt tca aaa atg gaa

61/21                                               91/31
caa aga tca ctc ctt tgc ttc ctt tat ctg ctc cta cta ttc aat ttc act ctc aga gtc

121/41                                              151/51
gct gga aac gct gaa ggt gat gct ttg act cag ctg aaa aac agt ttg tca tca ggt gac

181/61                                              211/71
cct gca aac aat gta ctc caa agc tgg gat gct act ctt gtt act cca tgt act tgg ttt

241/81                                              271/91
cat gtt act tgc aat cct gag aat aaa gtt act cgt gtg gag ctt tat agc aat aac att

301/101                                             331/111
aca ggg gag ata cct gag gag ctt ggc gac ttg gtg gaa cta gta agc ttg gat ctt tac

361/121                                             391/131
gca aac agc ata agc ggt ccc atc cct tcg tct ctt ggc aaa cta gga aaa ctc cgg ttc

421/141                                             451/151
ttg cgt ctt aac aac aat agc tta tca ggg gaa att cca atg act ttg act tct gtg cag

481/161                                             511/171
ctg caa gtt ctg gat atc tca aac aat cgg ctc agt gga gat att cct gtt aat ggt tct

541/181                                             571/191
ttt tcg ctc ttc act cct atc agt ttt gcg aat aat agc tta acg gat ctt ccc gaa cct

601/201                                             631/211
ccg cct act tct acc tct cct acg cca cca cca cct tca ggg ggg caa atg act gca gca

661/221                                             691/231
ata gca ggg gga gtt gct gca ggt gca gca ctt cta ttt gct gtt cca gcc att gcg ttt

721/241                                             751/251
gct tgg tgg ctc aga aga aaa cca cag gac cac ttt ttt gat gta cct gct gaa gaa gac

781/261                                             811/271
cca gag gtt cat tta gga caa ctc aaa agg ttt acc ttg cgt gaa ctg tta gtt gct act

841/281                                             871/291
gat aac ttt agc aat aaa aat gta ttg ggt aga ggt ggt ttt ggt aaa gtg tat aaa gga

901/301                                             931/311
cgt tta gcc gat ggc aat cta gtg gct gtc aaa agg cta aaa gaa gaa cgt acc aag ggt

961/321                                             991/331
ggg gaa ctg cag ttt caa acc gaa gtt gag atg atc agt atg gcc gtt cat agg aac ttg

1021/341                                            1051/351
ctt cgg ctt cgt ggc ttt tgc atg act cca act gaa aga tta ctt gtt tat ccc tac atg

1081/361                                            1111/371
gct aat gga agt gtt gct tct tgt tta aga gag cgt cct gaa ggc aat cca gca ctt gat

1141/381                                            1171/391
tgg cca aaa aga aag cat att gct ctg gga tca gca agg ggg ctt gcg tat tta cat gat

1201/401                                            1231/411
cat tgc gac caa aaa atc att cac cgg gat gtt aaa gct gct aat ata ttg tta gat gaa

1261/421                                            1291/431
gag ttt gaa gct gtt gtt gga gat ttt ggg ctc gca aaa tta atg aat tat aat gac tcc

1321/441                                            1351/451
cat gtg aca act gct gta cgc ggt aca att ggc cat ata gcg ccc gag tac ctc tcg aca

1381/461                                            1411/471
gga aaa tct tct gag aag act gat gtt ttt ggg tac ggg gtc atg ctt ctc gag ctc atc
```

FIGUUR 18a CONTD.

1441/481 1471/491
act gga caa aag gct ttc gat ctt gct cgg ctt gca aat gat gat gat atc atg tta ctc

1501/501 1531/511
gac tgg gtg aaa gag gtt ttg aaa gag aag aag ttg gaa agc ctt gtg gat gca gaa ctc

1561/521 1591/531
gaa gga aag tac gtg gaa aca gaa gtg gag cag ctg ata caa atg gct ctg ctc tgc act

1621/541 1651/551
caa agt tct gca atg gaa cgt cca aag atg tca gaa gta gtg aga atg ctg gaa gga gat

1681/561 1711/571
ggt tta gct gag aga tgg gaa gaa tgg caa aag gag gag atg cca ata cat gat ttt aac

1741/581 1771/591
tat caa gcc tat cct cat gct ggc act gac tgg ctc atc ccc tat tcc aat tcc ctt atc

1801/601 1831/611
gaa aac gat tac ccc tcg ggt cca aga taa cct ttt aga aag ggt ctt ttc ttg tgg gtt

1861/621
ctt caa caa gta tat ata tag att ggt gaa gtt tta aga tgc aaa aaa aa

Figure 18b

Predicted amino acid sequence of the Arabidopsis thaliana RKS-13 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 4 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of
4 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

ME[illegible]YLYLL
[illegible]ACNAEG

DA[illegible]SSGDP
A[illegible]ATLVT

[illegible]NPLNKVTRV

ELYSNNITGEI
P[illegible]VSLDLYANSISGPI
P[illegible]RFLRLNNNSLSGEI
P[illegible]VLDISNNRLSGDI
P[illegible]TPISFANNSLTDLPE

P[illegible]PTPPPPSG

[illegible]VAAGAAL
[illegible]AAWL

[illegible]VPGAEEDPE
[illegible]PTLRELLVAT

[illegible]VLGRGGFGKVYKGRLAD
[illegible]LKEERTKGGELQFQ
[illegible]AVHRNLLRLRGFCM
[illegible]PYMANGSVASCLR
[illegible]PKRKHIALGSA
[illegible]CDQKIIHRDVKAA
[illegible]AVVGDFGLAKLMN
Y[illegible]AVRGTIGHIAPEYL
ST[illegible]DVFGYGVMLLELI
[illegible]ANDDDIMLLDW
V[illegible]LISLVDAELEGKY
V[illegible]ALLCTQSSAME
[illegible]

[illegible]KEEMPIHDFNYQA

Y[illegible]YSNSLIENDYPSGPR

Figure 19a
Arabidopsis thaliana RKS14 cDNA
The start codon has been indicated by bold capitals.

```
1/1                                     31/11
ctg cac ctt aga gat taa tac tct caa gaa aaa caa gtt ttg att cgg aca aag ATG ttg

61/21                                   91/31
caa gga aga aga gaa gca aaa aag agt tat gct ttg ttc tct tca act ttc ttc ttc ttc

121/41                                  151/51
ttt atc tgt ttt ctt tct tct tct tct gca gaa ctc aca gac aaa gtt gtt gcc tta ata

181/61                                  211/71
gga atc aaa agc tca ctg act gat cct cat gga gtt cta atg aat tgg gat gac aca gca

241/81                                  271/91
gtt gat cca tgt agc tgg aac atg atc act tgt tct gat ggt ttt gtc ata agg cta tac

301/101                                 331/111
agg tta ttg cag aac aat tac ata aca gga aac atc cct cat gag att ggg aaa ttg atg

361/121                                 391/131
aaa ctc aaa aca ctt gat ctc tct acc aat aac ttc act ggt caa atc cca ttc act ctt

421/141                                 451/151
tct tac tcc aaa aat ctt cac agg agg gtt aat aat aac agc ctg aca gga aca att cct

481/161                                 511/171
agc tca ttg gca aac atg acc caa ctc act ttt ttg gat ttg tcg tat aat aac ttg agt

541/181                                 571/191
gga cca gtt cca aga tca ctt gcc aaa aca ttc aat gtt atg ggc aat tct cag att tgt

601/201                                 631/211
cca aca gga act gag aaa gac tgt aat ggg act cag cct aag cca atg tca atc acc ttg

661/221                                 691/231
aac agt tct caa aga act aaa aac cgg aaa atc gcg gta gtc ttc ggt gta agc ttg aca

721/241                                 751/251
tgt gtt tgc ttg ttg atc att ggc ttt ggt ttt ctt ctt tgg tgg aga aga aga cat aac

781/261                                 811/271
aaa caa gta tta ttc ttt gac att aat gag caa aac aag gaa gaa atg tgt cta ggg aat

841/281                                 871/291
cta agg agg ttt aat ttc aaa gaa ctt caa tcc gca act agt aac ttc agc agc aag aat

901/301                                 931/311
ctg gtc gga aaa gga ggg ttt gga aat gtg tat aaa ggt tgt ctt cat gat gga agt atc

961/321                                 991/331
atc gcg gtg aag aga tta aag gat ata aac aat ggt ggt gga gag gtt cag ttt cag aca

1021/341                                1051/351
gag ctt gaa atg ata agc ctt gcc gtc cac cgg aat ctc ctc cgc tta tac ggt ttc tgt

1081/361                                1111/371
act act tcc tct gaa cgg ctt ctc gtt tat cct tac atg tcc aat ggc agt gtc gct tct

1141/381                                1171/391
cgt ctc aaa gct aaa ccg gta ttg gat tgg ggc aca aga aag cga ata gca tta gga gca

1201/401                                1231/411
gga aga ggg ttg ctg tat ttg cat gag caa tgt gat cca aag atc att cac cgt gat gtc

1261/421                                1291/431
aaa gct gcg aac ata ctt ctt gac gat tac ttt gaa gct gtt gtc gga gat ttc ggg ttg

1321/441                                1351/451
gct aag ctt ttg gat cat gag gag tcg cat gtg aca acc gcc gtg aga gga aca gtg ggt

1381/461                                1411/471
```

FIGUUR 19a CONTD.

cac att gca cct gag tat ctc tca aca gga caa tct tct gag aag aca gat gtg ttc ggt

1441/481 1471/491
ttc ggg att ctt ctt ctc gaa ttg att act gga ttg aga gct ctt gaa ttc gga aaa gca

1501/501 1531/511
gca aac caa aga gga gcg ata ctt gat tgg gta aag aaa cta caa caa gag aag aag cta

1561/521 1591/531
gaa cag ata gta gac aag gat ttg aag agc aac tac gat aga ata gaa gtg gaa gaa atg

1621/541 1651/551
gtt caa gtg gct ttg ctt tgt aca cag tat ctt ccc att cac cgt cct aag atg tct gaa

1681/561 1711/571
gtt gtg aga atg ctt gaa ggc gat ggt ctt gtt gag aaa tgg gaa gct tct tct cag aga

1741/581 1771/591
gca gaa acc aat aga agt tac agt aaa cct aac gag ttt tct tcc tct gaa cgt tat tcg

1801/601 1831/611
gat ctt aca gat gat tcc tcg gtg ctg gtt caa gcc atg gag tta tca ggt cca aga tga

1861/621 1891/631
caa gag aaa cta tat gaa tgg ctt tgg gtt tgt aaa aaa

Figure 19b
Predicted amino acid sequence of the Arabidopsis thaliana RKS-14 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as described in Schmidt et al. (1997).
At the predicted extracellular domain the first domain represents a signal sequence.
The second domain contains a leucine zipper motif, containing 2 leucine residues, each separated by 7 other amino acids.
The third domain contains conserved cysteine residues, involved in disulphate bridge formation.
The fourth domain contains a leucine rich repeat domain, consisting of 4 complete repeats of each approximately 24 amino acid residues.
The fifth domain contains many serine and proline residues, and is likely to contain hydroxy-proline residues, and to be a site for O-glycosylation.
The sixth domain contains a single transmembrane domain after which the predicted intracellular domains are positioned.
The seventh domain has an unknown function.
The eight domain represents a serine/threonine protein kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions.
The ninth domain has an unknown function.
The last and tenth domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

MLQGRREAKKSYALFSSTFF
FFFICFLSSSSAELTDKV

VALIGIKSSLTDP
HGVLMNWDDTAVD

PCSWNMITCSDGFVIR

LYRLLQNNYITGNI
PHEIGKLMKLKTLDLSTNNFTGQI
PFTLSYSKNLHRRVNNNSLTGTI
PSSLANMTQLTFLLDLSYNNLSGPV
PRSLA KTFNVMGNSQICPT

GTEKDCNGTQPKPMSITLNSSQRGTKNRK

IAVVFGVSLTCVCLLIIGFGFLLWW

RRRHNKQVLFFDINEQNKE
EMCLGNLRRFNFKELQSAT

SNFSSKNLVGKGGFGNVYKGCLHD
GSIIAVKRLKDINNGGGEVQFQ
TELEMISLAVHRNLLRLYGFCT
TSSERLLVYPYMSNGSVA
SRLKAKPVLDWGTRKRIALGAG
RGLLYLHEQCDPKIIHRDVKAA
NILLDDYFEAVVGDFGLAKLLD
HEESHVTTAVRGTVGHIAPEYL
STGQSSEKTDVFGFGILLLELI
TGLRALEFGKAANQRGAILDW
VKKLQQEKKLEQIVDKDLKSNY
DRIEVEEMVQVALLCTQYLPIH
RPKMSEVVRMLE

GDGLVEKWEASSQRAET
NRSYSKPNEFSSS

ERYSDLTDDSSVLVQAMELSGPR

Figure 20 A
Arabidopsis thaliana RKS 7 partial cDNA sequence.
The 5'-end and a region between the two cDNA fragments (.....) is not shown.

AGCGAATATACTTCTTGATGACTACTGTGAAGCTGTGGTTGGCGATTTTGG
TTTAGCTAAACTCTTGGATCATCAAGATTCTCATGTGACAACCGCGGTTAG
AGGCACGGTGGGTCACATTGCTCCAGAGTATCTCTCAACTGGTCAATCCTC
T.....
AACAGATGTTTTTTGGCTTTGGGATTCTTCTTCTTGAGCTTGTAACCGGAC
AAGGAGCTTTTGAGTCTGTTAAAGCGGCTAACCGGAAAGGTGTGATGCTTG
ATTGGGTTAAAAAGATTCATCAAGAGAAGAAACTTGAGCTACTTGTGGATA
AAGAGTTGTTGAAGAAGAAGAGCTACGATGAGATTGAGTTAGACGAAATGG
TAAGAGTAGCTTTGTTGTGCACACAGTACCTGCCAGGACATAGACCAAAAA
TGTCTGAAGTTGTTCGAATGCTGGAAGGAGATGGACTTGCAGAGAAATGGG
AAGCTTCTCAAAGATCAGACAGTGTTTCAAAATGTAGCAACAGGATAAATG
AATTGATGTCATCTTCAGACAGATACTCTGATCTTACCGATGACTCTAGTT
TACTTGTGCAAGCAATGGAGCTCTCTGGTCCTAGATGAAATCTATACATGA
ATCTGAAGAAGAAGAAGAACATGCATCTGTTTCTTGAATCAAGAGGGATTC
TTGTTTTTTTGTATAATAGAGAGGTTTTTTGGAGGGAAATGTTGTGTCTCT
GTAACTGTATAGGCTTGTTGTGTAAGAAGTTATTACTGCACTTAGGGTTAA
TTCAAAGTTCTTTACATAAAAAATGATTAGTTGCGTTGAATAGAGGGAACA
CTTTGGGAGATTTCATGTATGAAATTTGG

Figure 20B
Predicted partial amino acid sequences of the Arabidopsis thaliana RKS-7 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

A
NILLDDYCEAVVGDFGLAKLLD
HQDSHVTTAVRGTVGHIAPEYL
STGQSS..QMFFGFGILLLELV
TGQGAFE SVKAANRKGVMLDW
VKKIHQEKKLELLVDKELLKKKSY
DEIELDEMVRVALLCTQYLPGH
RPKMS EVVRMLE

GDGLAEKWEASQRSDS
VSKCSNRINELMSSS

DRYSDLTDDSSLLVQAMELSGPR*

Figure 21 A
Arabidopsis thaliana RKS 9 partial cDNA sequence.
The 5'-end is not shown.

GAAATGGTAAGAGTAGCTTTGTTGTGCACACAGTACCTGCCAGGACATAGA
CCAAGAGTGTCTGAAGTTGTTCGAATGCTGGAAGGAGATGGACTTGCAGAG
AAGTGGGAAGCTTCTCAAGGATCAGACAGTGTTTCAAAATGTAGCAACAG
GATAAATGAAGTGATGTCATCTTCAGACAGATACTCTGATGTTACCGATGA
CTCTAGTTTACGTGTGCAAGCAATGGAGCTCTCTGGTCCTAGATGAAGTCT
ATACATGAATCTGAAGAAGAAGAAGAACATGCATCTGTTTCTTGAATCAAG
AGGGATTCTTGTTTTTTTGTATAATAGAGAGGTTTTTTGGAGGGAAATGTT
GTGTCTCTGTAACTGTATAGGCTTGTTGTGTAAGAAGTTATTACTGCACTT
AGGGTTAAGTCAAAGTTCTTTACATAAGGGGGGATTAGTTGCGTTGAATAG
AGGGAACACTTTGGGAGATTTCATGTGTGAAAGTTGGGAAGTCATGTTTGA
GAATGAAGGTTATCTTATTATTGAA

Figure 21B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-9 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

VDKELLKKKSY
DEIELDEMVRVALLCTQYLPGH
RPRVSEVVRMLE

GDGLAEKWEASQGSDS
VSKCSNRINEVMSSS

DRYSDVTDDSSLRVQAMELSGPR*

Figure 22A

Arabidopsis thaliana RKS 15 partial cDNA sequence.
The 5'-end is not shown.

GTGGATAAAGAGTTGTTGAAGAAGAAGAGCTACGATGAGATTGAGTTAGA
CGAAATGGTAAGAGTAGCTTTGTTGTGCACACAGTACCTGCCAGGACATA
GACCAAGAGTGTCTGAAGTTGTTCGAATGCTGGAAGGAGATGGACTTGCA
GAGAAGTGGGGAAGCTTCTCAAGGATCAGACAGTGTTTCAAAATGTAGCA
ACAGGATAAATGAAGTGATGTCATCTTCAGACAGATACTCTGATGTTACC
GATGACTCTAGTTTACGTGTGCAAGCAATGGAGCTCTCTGGTCCTAGATG
AAGTCTATACATGAATCTGAAGAAGAAGAAGAACATGCATCTGTTTCTTG
AATCAAGAGGGATTCTTGTTTTTTTTGTATAATAGAGAGGTTTTTTGGAGG
GAAATGTTGTGTCTCTGTAACTGTATAGGCTTGTTGTGTAAGAAGTTATT
ACTGCACTTAGGGTTAAGTCAAAGTTCTTTACATAAGGGGGGATTAGTTG
CGTTGAATAGAGGGAACACTTTGGGAGATTTCATGTGTGAAAGTTGGGAA
GTCATGTTTGAGAATGAAGGTTATCTTATTATTGAA

Figure 22B

Predicted amino acid sequence of the Arabidopsis thaliana RKS-15 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

VDKELLKKKSY
KEIELDEMVRVALLCTQYLPGH
RPRVSEVVRMLE

GDGLAEKWEASQGSDSVSKCSNRINEVMSSS

DRYSDVTDDSSLRVQAMELSGPR*

Figure 23A
Arabidopsis thaliana RKS 16 partial cDNA sequence.
The 5'-end is not shown.

AAAGTACGTGGAAGCAGAAGTGGAGCAGCTGATACGAATGGCTCTGCTCTG
CACTCAAAGTTCTGCAATGGAACGTCCAAAGATGTCAGAAGTAGTGAGAAT
GCTGGAAGGAGATGGTTTAGCTGAGAGATGGGAAGAATGGCAAAAGGAGGA
GATGCCAATACATGATTTTAACTATCAAGCCTATCCTCATGCTGGCACTGA
CTGGCTCATCCCCTATTCCAAGTCCCTTATCGAAGGCGATTACCCCTCGGG
TCCAAGATAACCTTTTAGAAAGGGTCTTTTCTTGTGGGTTCTTCAACAAGT
ATATATATAGATTGGTGAAGTTTTAAGATGCAAGAGGGGGCCATGCACTTT
TGAATATCACCTCCTCTATAAGTAGTATTGTGTCTCTTG

Figure 23B
Predicted amino acid sequence of the Arabidopsis thaliana RKS-16 protein. Different domains are spaced and shown from the N-terminus towards the C-terminus. Overall domain structure is similar as descibed in Schmidt et al. (1997). The protein sequence is obtained from partial cDNA sequences. The first available domain represents part of a serine/threonine protien kinase domain (Schmidt et al. 1997), and is probably also containing sequences for protein, protein interactions. The next domain has an unknown function. The last domain at the C-terminal end represents a single leucine rich repeat, probably involved in protein, protein interactions.

KY
VEAEVEQLIRMALLCTQSSAME
RPKMSEVVRMLE

GDGLAERWEEWQKEEMPIHDFNYQAY

PHAGTDWLIPYSKSLIEGDYPSGPR*